

## **DOCTORAL THESIS**

### **Inaction and Silent Action in Interaction**

Berger, Israel

*Award date:*  
2013

*Awarding institution:*  
University of Roehampton

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Inaction and Silent Action in Interaction

by

Israel Berger, BA(Hons.) cum laude

A thesis submitted in partial fulfilment of the degree of Doctor of Philosophy

Department of Psychology

University of Roehampton

2012

### **Abstract**

How do non-vocal practices function within sequences? This thesis addresses silence and gesture in the context of social interaction involving human participants through the framework of conversation analysis (CA) to answer this question. Although it is certainly possible, and indeed common, for gestures and other non-vocal practices to occur during talk, this thesis focuses on those that occur without accompanying talk. In order to understand the role of non-vocal practices in this environment, we must first understand the role of silence (or the absence of talk) in participants' interactions. How does silence function within the sequential environment? How does context affect how silence and non-vocal practices are treated by participants? If one organisation (or aspect of an organisation) is affected, are all organisations (or aspects of that organisation) affected? I draw on psychological, sociological, and linguistic literature to show why silence and gesture are related in interactional research and how this affects conversation analytic methodology. The work that forms this thesis brings together cross-cultural perspectives and technical advances with respect to silence and non-vocal practices both individually and when they occur together (i.e. non-vocal practices without accompanying talk). I begin with a broad overview of research and theories of gesture and silence before discussing CA as a method and its relationship to silence and non-vocal practices.

The empirical studies begin with silence in relation to culture and context and issues in analysing and transcribing silence. I then examine how one sequential environment, in which psychotherapy clients are obligated to respond by orienting to the therapeutic agenda, has a preference structure that is very different from ordinary conversation. In this sequential environment, silence and elaboration are marks of preferred responses rather than dispreferred. The preference structure is made particularly visible through accountability that becomes relevant when a client's response is produced promptly following the therapist's overtly therapeutic action. Silence in this environment contributes

to clients' performance of sincerity and participation in the psychotherapeutic process, and non-vocal practices during longer silences can show that the client remains engaged with the sequence. Many authors have accounted for silences that are not treated by participants as problematic by applying constructs such as 'continuing states of incipient talk'. This construct, however, is variably used and has not been developed through empirical examination. It does not adequately explain interactions that involve silence or gesture, as I show through a content analysis and systematic review. After recommending that researchers engage with participant orientations in environments that differ from canonical conversations, I describe an environment that is commonly thought of as constituting a 'continuing state of incipient talk' – television-watching. I show that contrary to some claims about 'incipient talk' environments, although response relevance is relaxed, both sequence organisation and turn-taking are strongly oriented to by participants. Compared to ordinary conversation, television-watching also involves more gestures and other non-vocal practices without accompanying talk. I examine how non-vocal practices without accompanying talk are used in interaction. As responsive actions, gestures can be used without accompanying talk as a resource for doing sensitive interactional work, particularly in places where giving offence might be a concern. Non-vocal practices can also be used in other situations to accomplish sequential actions that could otherwise be spoken. These uses of non-vocal practices create methodological questions for conversation analysis, which has traditionally focused on the *talk* of participants. By clearly distinguishing between *actions* and *turns* (two classic CA concepts) and examining the timing of non-vocal practices, I show that non-vocal practices can have a clearly defined role in sequence organisation. CA can thus be a useful method for examining the entire situation of social interaction, including non-vocal practices.

## **Contents**

### **Acknowledgements**

## **1 Introduction: Silence and Gesture in Interaction and Their Relationship to**

### **Conversation Analysis**

#### 1 Introduction

##### 1.1 What is Non-vocal?

#### 2 Silence in Interaction

##### 2.1 Contextual Considerations

##### 2.2 Conversation Analysis and Silence

#### 3 Gesture Research

##### 3.1 Early Gesture Research

##### 3.2 Interactional Research

##### 3.3 Conversation Analysis and Gesture

#### 4 Conclusion

## **2 An Overview of Conversation Analysis and the Data Presented**

#### 1 Introduction

#### 2 History of Conversation Analysis

#### 3 For What Research Questions is Conversation Analysis Useful?

#### 4 Context of the Thesis

#### 5 The Data

##### 5.1 Coral & Al

##### 5.2 Paula & Leif

##### 5.3 Paula & Sally

5.4 Kim & Barbara, Dana & Niles

6 Application of Conversation Analysis to Psychotherapy and Related Fields

7 Doing Conversation Analysis

7.1 Cultural Intelligence in Doing Conversation Analysis

7.2 Conveying Information

7.3 Structures of Social Interaction

7.3.1 Sequential Organisation (Overall Structural Organisation)

7.3.2 Sequence Organisation

7.3.3 Action Formation

7.3.4 Turn Taking

7.3.5 Preference

7.3.6 Trouble/Repair

7.3.7 Word Selection

7.3.8 Topic

8 Conversation Analysis with Silence and Non-vocal Practices

9 Rationale for Using Conversation Analysis for this Project

**3 Support and Evidence for Considering Local Contingencies in Studying and**

**Transcribing Silence in Conversation**

1 Introduction

2 Deviations and Support

2.1 Deviation from Canonical Turn Taking Features

2.2 Orientation to the Larger Organisations of Turns and Sequences

3 Considerations for Transcription

3.1 Giving Approximate Values to Silences by Using a Fixed Rating  
System

3.2 Using a Conversation-Based Metric for Timing Silences

3.3 Taking the Mean Silence Between Words During Smoothly Flowing  
Talk and Subtracting This from Timed Silences

3.4 Timing Silences as They Are and Conveying in a Description of the  
Material that Longer than Canonical Silences are Typical of the  
Interaction

4 Conclusion

#### **4 How Silence Contributes to the Performance of Sincerity in Psychotherapy**

1 Introduction

2 The Data

3 Immediacy is Dispreferred: Accounting for Immediate Responses to Therapeutic  
Questions

3.1 Structurally Provided-For Responses

3.2 Immediate Responses to 'Obvious' Questions

3.3 Disaffiliative Immediate Responses

4 Silence is Preferred: An In-Depth Analysis

4.1 Characteristic Delayed 'Sincere' or 'Thoughtful' Responses

4.2 Uses of Quick Responses

4.1.1 Quick Responses that Minimise

4.1.2 Quick Responses that Challenge

4.3 Re-establishing Alignment and Affiliation

4.4 Increments and Reformulated Turns versus Continuation

5 Conclusion

## **5 Do ‘Continuing States of Incipient Talk’ Exist?**

### 1 Introduction

### 2 Content Analysis

#### 2.1 Variables

#### 2.2 Results

#### 2.3 Discussion

### 3 Silence and Sequences

#### 3.1 Pauses

#### 3.2 Gaps

#### 3.3 Lapses

### 4 Unpacking Incipient Talk

### 5 Characterisations of Incipient Talk

#### 5.1 Lack of Openings and/or Closings

#### 5.2 Lack of Conditional Relevance

#### 5.3 Silences are Common

#### 5.4 Lapses During Talk-Based Interaction

#### 5.5 Multi-Party Engagement

#### 5.6 Summons-Answer Sequences

### 6 Some Possibly Related Concepts

#### 6.1 Unfocused Interactions

#### 6.2 Open States of Talk

#### 6.3 Gesprächsinseln

#### 6.4 Confusion and Conflation

### 7 Do ‘Continuing States of Incipient Talk’ Exist?



## **6 How Television Plays into Sequence Organisation: Orientation to Conversational Structures During Television-Watching**

### 1 Introduction

### 2 Incipient Talk

### 3 Sequence Organisation and Turn-taking

#### 3.1 Adjacency Pairs

#### 3.2 Speaker Selection and Turn Allocation

### 4 Orientation to Existing Structures

#### 4.1 Adjacency Pairs

#### 4.2 Minimal Responses

#### 4.3 Preference

### 5 Television-Specific Orientation: Television as a Topic Resource

### 6 Actions

#### 6.1 Triggers for Initiating Actions

#### 6.2 Responding to Initiating Actions

#### 6.3 Pursuit

### 7 Accounting for the Differences

#### 7.1 Interactional Focus

#### 7.2 Watching as an Event

#### 7.3 Sensory Limitations

#### 7.4 Acceptability and Sanctioning

### 8 Conclusion

## **7 Some Uses of Gestural Responsive Actions**

### 1 Introduction

### 2 Concurrent Gesture

### 3 Responsive Gestures

#### 3.1 Sequence Organisation and Gesture

### 4 Uses for Gestural Responses

#### 4.1 Sequentially Positioned Displays of Stance

#### 4.2 Use of One Action Type to Do Another

#### 4.3 A Possible Use: Bracketing Intervening Talk

#### 4.4 Using Gestures to Show that Talk is Forthcoming

### 5 Repair on Gesture

### 6 Conclusion

## **8 Are Non-vocally Implemented Actions Subject to the Same Organizational Constraints as Talk-implemented Actions?**

### 1 Non-Vocal Behaviour as an Issue in the Sequential Analysis of Conversation

### 2 Action Formation

### 3 Practices that Occur in Structurally Sequential Positions

#### 3.1 Terminology: Describing Behaviour

#### 3.2 Communicative Non-vocal Practices

#### 3.3 Tasks

### 4 Action Timing: Onset, Duration and Transition

#### 4.1 Action Onset

#### 4.2 Permeability of Non-Vocal Practices

### 5 When Non-Vocal Practices Constitute a Part of the Overall Action

### 6 When Non-Vocal Practices Constitute *the* Action

#### 6.1 As Responsive Actions

#### 6.2 As Initiating Actions

### 7 Conclusion

## **9 Conclusion: Developing a Comprehensive Conversation Analytic Methodology**

1 Introduction

2 A Critical Evaluation

3 Original Contributions

4 Implications and Suggestions for Future Research

## **References**

## **Appendices**

A. Transcription Key

B. CA Data Bank Consent Form

C. Good Relations Ethical Approval

D. Good Relations Recruitment Letters

E. Good Relations Information Sheet

F. Good Relations Consent Form

G. Good Relations Debriefing Letter

H. Consent Form for KB and DN Clips

I. Permission to use KB and DN Clips

J. Chapter 5 Bibliography of Content Analysis Sample

K. Chapter 7 CA052109 'Cooked it too fast'

### **Acknowledgements**

First and foremost, I would like to thank my family for their continued and unconditional support through my studies and personal development. The last thing my grandfather wanted me to know was that I could do anything, and this stays with me. My mother and father encouraged me to learn according to my curiosity and provided me with the tools to do so. My mother and step-father have shown great dedication to my postgraduate studies and have supported me through difficulties. My husband Y. Gavriel Ansara has gone through the struggles of a PhD in psychology alongside me, and his professional, intellectual and emotional support have kept me going and improved my work. Philip Kwok carried me through difficulties performing statistical tests that were new to me, in particular explaining potential shortfalls in cluster analysis and how to prevent them. Other people who have been there for me are too numerous to name individually, and I would likely inadvertently exclude some. This work would surely not have been attempted without all of them. This work would also not have been possible without participants' support of the research from which this thesis draws and from which I have had the honour to publish results elsewhere. I am similarly grateful to Terhi Korkiakangas and Rowena Viney for allowing me to use portions of their data to assist in illustrating some of the analyses.

Although the text of chapters five and six was written by me, discussion, data sessions, and comments by Rowena Viney were essential for the development of these chapters. She was also a second coder for the content analysis in chapter five. Discussion with Terhi Korkiakangas was helpful in developing the arguments in chapters seven and eight, and her friendship and advice has helped me through the inevitable challenges of completing a PhD. As my first year mentor, Jon Silas helped me feel at home and remains a valuable friend and colleague. My Director of Studies, John Rae, continuously exceeded

the call of duty in his enthusiasm and feedback. No blurb in an acknowledgements section does his contributions to my postgraduate education and professional development justice. I am most grateful for the professional contacts I have made through him, the opportunities he has made available to me through those introductions, and the insightful feedback on my work that has led to the improvement of this and other works. Thank you to Mick Finlay for introducing me to John and for encouraging me. Thank you to my Co-Supervisor, Lyndsey Moon, for her support and especially for not being afraid to stand up for me in the face of bureaucracy. Finally, thanks are due to Sue Speer for introducing me to conversation analysis and authoritative methodological texts.

Many of the chapters of this thesis have been presented at conferences, published, or submitted for publication. I am particularly indebted to colleagues such as Jürgen Streeck, who was the panel discussant for the presentation that became chapters 7 and 8 at the International Conference for Conversation Analysis in 2010 in Mannheim, where the presentation that evolved into chapters 5 and 6 was given. Versions of these chapters were also presented at the Conference of the International Society for Gesture Studies 2010 in Frankfurt (Oder), and my participation at this conference was supported by a bursary from the International Society of Gesture Studies. Versions of Chapters 1 (*Silence and Gesture in Interaction: Their Relationship to Interactional Linguistics and Conversation Analysis*), 4 (*How Silence Contributes to the Performance of Sincerity in Psychotherapy*), 5 (*Do ‘Continuing States of Incipient Talk’ Exist?*), and 6 (*How Television Plays into Sequence Organisation: Orientation to Conversational Structures During Television-Watching*) are under review for publication. A version of Chapter 5 was additionally presented at the Discourse-Communication-Conversation conference at Loughborough University in 2012, for which I received a conference bursary from the Psychology Postgraduate Affairs Group (PsyPAG). Thank you to Rein Sikveland for assisting with the translation of relevant

paragraphs of the Swedish article in the content analysis data set. Chapter 8 (*Are Non-vocally Implemented Actions Subject to the Same Organizational Constraints as Talk-implemented Actions?*) is currently under revision at a peer reviewed journal. A version of Chapter 3 (*Support and Evidence for Considering Local Contingencies in Studying and Transcribing Silence in Conversation*) has been published as: Berger, I. (2011). Support and evidence for considering local contingencies in studying and transcribing silence in conversation. *Pragmatics*, 21(3), 291-306. A version of Chapter 7 (*Some Uses of Gestural Responsive Actions*) has been published as: Berger, I. & Rae, J. (2012). Some uses of gestural responsive actions. *Journal of Pragmatics*, 44(13), 1821-1835. Thanks are due to all of the people I have mentioned and any I have not enumerated, however any errors or deficits remain mine.

## **Chapter 1**

### **Introduction: Silence and Gesture in Interaction and Their Relationship to Conversation Analysis<sup>1</sup>**

---

<sup>1</sup> A version of this chapter has been submitted to a peer-reviewed journal and is awaiting review.

### **Abstract**

In this chapter, I provide a brief, historical account of key research and theories concerning silence and gesture. In the case of gesture research, the focus is on those that occur without accompanying talk, however the nature of gesture research makes consideration of concurrent gestures necessary. After establishing the current state of silence and gesture research, particularly with relevance to conversation analytic methodology, I discuss the challenge that non-vocal practices such as gestures pose for *sequence organisation*. Although the internal sequentiality of non-vocal practices (e.g., trajectory) has been an object of study with broadly applicable findings, little research has addressed the role of non-vocal practices within other sequential structures largely organised by talk.



## **[1] Introduction**

Human behaviour involves a variety of practices that can be categorised as silence or sound, and the interpretation of sounds involving the voice has been the primary aim of linguistics and psychology of language. Non-vocal practices (themselves involving silence) such as gesture can occur during instances of vocal sound or silence. Silence and gesture have held important places in the social sciences and have been the subject of much research and debate. Of particular interest has been the extent to which differences and/or similarities exist across cultures and situations (e.g., Matsumoto & Willingham, 2006; Ekman & Keltner, 1997; Stivers et al., 2009; Gardner, Fitzgerald, & Mushin, 2009) and how observers can discern the ‘true’ intentions or emotions of people (e.g., Ekman & Friesen, 1969a; Matsumoto & Willingham, 2006; Scherer, Feldstein, Bond, & Rosenthal, 1985). Inter-species similarities in gesture use among primates has been growing as a subject of study over the past 20 years (e.g., Preuschoft & van Hooff, 1997; Suomi, 1997; Rossano, 2011) and is foundational to some areas of evolutionary psychology. Although this thesis will occasionally draw from studies that examine the communicative behaviours of non-human primates either solely or alongside human primates, its objective is to examine the interactional relevance of silence and non-vocal behaviour without accompanying utterances in human interactions. I will primarily draw from interactional research in order to maintain a focus on how people actually use the interactional resources that are available to them – rather than how authors propose resources ought to be used or how people in laboratory settings perform or interpret practices.

Human behaviour develops and is governed by culture and situation, and issues such as personality or demographic statuses can become relevant in interaction as well. Before proceeding, a distinction between broad, cultural and specific, situational contexts must be made. By culture, most authors refer to participants’ ethnic backgrounds and

countries of origin and by situation, the particular institutions or tasks in which participants are involved. However, as the strong traditions of ethnography and ethnomethodology that have developed over the past approximately 60 years has conveyed, cultures can differ between institutions or families as widely as they can between countries or ethnic groups. As Schefflen (1974, p. 97) notes, observing a person's behaviour can help us identify their 'origins and institutional experience', which are not unidimensional. Two or more cultures can be present in the same situation for the same people. For example, interactions within a law firm in Uruguay would integrate both mainstream Uruguayan culture and mainstream legal culture in addition to possibly other cultures as well, e.g., the religious affiliation(s) of staff, a broader South American culture, and immigrant ethnic culture(s) of staff. Beyond larger group cultures, the firm would have its own organisational culture, and individual teams would have their own (Millward, 2005). For the most part, however, interactional linguistic and social psychological research has viewed the construct of culture as a product of general location (or country) and ethnicity and the practices of more specific settings or family units as situational, specialised, or idiosyncratic.

It is unlikely that a given practice, whether performed with the voice or with the body, can be invented by one person and remain in use solely by that particular person. Without spreading, even if only to a family unit or group of co-workers, such a practice would not be communicative. Communicative gestures, for example, activate mirror neurons, which contribute to our social-behavioural learning (e.g., Rizzolatti & Craighero, 2004). It is thus also unlikely that a particular practice of which interlocutors demonstrate understanding has not been performed by someone else in some form, or such a practice would not be understood. The tendency to label practices as idiosyncratic reflects ignorance of other cultures and subcultures (Schefflen, 1974, p. 101). Rather, ambiguity and meaning work against one another to balance communicative flexibility and precision.

Scheflen (1974, p. 48) summarises the negative correlation of this relationship within Bateson's (1972) theoretical framework, 'meaning increases as ambiguity decreases, and ambiguity is decreased by the formation of larger and larger integrations of patterned behavior'. Implied in Scheflen's description, however, is that the relationship is also a progressive one. The idea that there might be benefits to maintaining ambiguity or flexibility is lost. However, communicative flexibility can be very important to conveying meaning in different sequential environments, as I will show in Chapter 7.

Gardner et al. (2009) also discuss how participants from the same country and ethnic background (Anglo-Australian) behave very differently in different interactional contexts (in this case, a political debate vs. relaxing at home). Gardner et al. thus reason that interactional differences are not 'cultural' per se but that an orientation to the goals of the interaction is pivotal in determining how interactants will behave. They propose that although differences in the extent to which particular types of practices (e.g., silence) may be seen across cultures, an underlying variable is how members of different cultures orient to the goals of both individual interactions and to social interaction more generally. In the case of silence, they propose that this is an orientation to, or lack of orientation to, pressure to 'get things done'. This argument is supported in Chapter 3 as well as by Lehtonen & Sajavaara (1985) who explain the origins of norms favouring silence in Finnish language and culture as possibly originating from the solitude of Finnish rural life. They go on to critique the American-centric approach to spoken interaction, in which silence is comparatively uncommon and is viewed as a lack of action, as avoidance of silence rather than genuine exchange of information (cf. Tannen, 1985 on New York vs. Californian communication styles).

East and Southeast Asian societies likewise tend to view talk as something that should be done with adequate forethought and for the purpose of exchanging information (Saville-Troike, 1985). However, even in Western European and North American discourses, silence among intimates is viewed as an indication of ‘interpersonal rapport so great that people understand each other without putting their thoughts into words’ (Tannen, 1985), which Baker (1955) terms ‘positive silence’. This is in contrast to Goffman’s (1967, p. 36) observation that undue lulls are signs of having nothing in common or of having to think of things to say (a form of what Baker (1955) terms ‘negative silence’), raising the question of what is undue and how much silence constitutes a lull. Baker (1955, p. 160) conceptualises silence in interaction as an association between tension and silence, with negative silence involving complete tension and positive silence involving no tension. He presents this as association, rather than movement along a continuum. All silences fall somewhere between fully positive and fully negative, and the same silence can *become* a different kind of silence. Baker’s system looks at the immediate interactional context of the silence, and subsequent silences can have different levels of tension.

As Sapir (1949, p. 53) observes, we often form judgements of people based on what they do and do not say. Within the study of silence, orientations to positive and negative politeness<sup>2</sup> (Brown & Levinson, 1987; see also Goffman, 1955) exist such that sometimes silence is seen as the avoidance of negative face (i.e. not performing an impolite action) or as the expression of positive face (i.e. communicating ‘good listening skills’). Even when silence is viewed as problematic, violations of positive and negative politeness are at work. By being silent, one may be neglecting to perform a polite action or communicating negative attitudes such as disinterest. Situationally, the meanings of silence and gesture can vary across cultures, and there are two main approaches to this issue. The

---

<sup>2</sup> Although Brown and Levinson’s (1987) theory and Goffman’s (1955) concept of face cannot be applied to all cultures (e.g., Matsumoto, 1988; Ide, 1989), they have a degree of usefulness in relation to our current discussion of silence.

cataloguing approach of descriptive linguistics (e.g., Morris, 1994) is particularly popular with gestures, whilst applied work problematising misunderstanding between members of different cultures (e.g., Eades, 2000, 2007) is equally applicable across silence, gesture, and other non-vocal behaviour. In this thesis, I will be particularly concerned with silence and gesture in sequences of action and contextual issues involved in their interpretation. I will also discuss how adequately addressing silence and gesture has methodological implications for conversation analysis (CA). I do not attempt to provide an exhaustive account of how silence or non-vocal behaviour is used in interaction but to address key issues in their analysis. In the following sections of this chapter, I will explore the background of these topics in terms of silence in interaction (Section 2) and gesture research (Section 3).

### *[1.1] What is Non-vocal?*

Saville-Troike (1985) distinguishes between verbal and non-verbal codes and vocal and non-vocal channels, as shown in Table 1. This is a useful starting point and certainly captures a wide range of communicative resources. However, for the purposes of the current analysis, a simplified system must suffice, as non-vocal verbal communication such as written language and sign languages are not within the scope of this thesis. Rather, I will use *verbal* to refer to only spoken language unless otherwise specified, and *vocal* will be used to refer to both spoken language and non-verbal vocalisations such as ‘mm’ or ‘ah’. *Non-vocal* will refer to any bodily behaviour that does not use the voice (gesture, performance of tasks, scratches, grooming, and the like). Because these behaviours are not necessarily communicative, although they may well communicate in a given context, it is necessary to further specify a set of terms to designate function. I use *behaviour* irrespective of ‘channel’ or function in social interaction, *practice* to refer to behaviours that are interactively consequential, and *resource* to refer to behaviours or classes of

behaviours that operate in a structurally regular way that can be drawn upon for effective communication.

Table 1. Codes and Channels, adapted from Saville-Troike's (1985)			
		Channel	
		Vocal	Non-vocal
Code	Verbal	Spoken language	Written language Sign languages Whistle/drum languages Morse code
	Non-verbal	Paralinguistic and prosodic features	Kinesics Proxemics Eye behaviour Pictures and cartoons

## [2] Silence in Interaction

In minority world<sup>3</sup> communication and communication theories, silence is commonly considered to be an abstaining, inaction, or withholding – an absence of communication (Scollon, 1985). Silence can be the background against which talk is perceived (Tannen, 1985), but it can also have communicative functions of its own, as Watzlawick, Beavin, and Jackson (1967, p. 48) argue: it is impossible for a person to not communicate. Some scholars take the position of silence as inaction a step further to ascribe personality characteristics to silence. Feldstein, Alberti, and BenDebba (1979, p. 85) group interactants

<sup>3</sup> *Minority world* refers to the approximately 17% of the world that is commonly called 'western' or 'developed'. This is in contrast to *majority world*, which refers to the approximately 83% of the world that is not generally included in these categories and may be called 'developing', 'underdeveloped', 'third world', etc. and also includes indigenous societies that are situated within geopolitical borders claimed by minority world states (Alam, 2007).

into two styles based on self-report questionnaires and recorded interactions: those who are ‘talkative, cheerful, and cooperative’ and those who are ‘reserved, detached, and taciturn’. They thus neglect the possibility that a person can be cheerful but quiet or talkative but detached. This binary view of silence in human communication styles is not the case worldwide (Kim, 2002, Chapter 13). More troubling, however, is that Feldstein et al. (1979) take these characteristics as causative of silence and describe people who are more prone to silence as ‘cold, suspicious, insecure, and tense’. Silence that occurs in the presence of two or more people (whether they are physically present or communicating through mediated systems) can mean nothing at all or it can mean more than one can communicate effectively using words. Silence inherently comes about as a result of *nobody* speaking and therefore *everybody* engaging in silence – what McDermott and Tylbor (1983) term *collusion*.

Despite collusion, silence in interaction is often attributable to a particular party or parties, making a verbal response relevantly absent (Sacks, Schegloff, & Jefferson, 1974). However, silence does not always simply mean absence. Although a response might be absent, the silence creates anticipation in parties who were not implicated in responding that can lead to additional, intervening talk or more attentive postures. Even when ‘nothing’ is happening, the implicated party can be *doing* doing nothing through their silence. Silence is *performative* and *communicative* in sometimes subtle ways. In co-present interaction, silence is inevitably accompanied by non-vocal behaviour, even if that behaviour is to remain still. By adopting a posture, one communicates ‘doing nothing’, ‘thinking’, etc. through one’s silence. Alternatively, interlocutors can produce a range of gestures with vague to precise meanings. All of these behaviours provide the context for others to understand the attributable silence, whereas silence on a telephone call might prompt one party to check whether the other is still on the line.

This is not to say that silence on the telephone does not convey meaning, only that one resource for avoiding conflict or saving face in such a situation is precisely the lack of additional information ordinarily provided in co-present interactions. That is, facial expressions and postures cannot ‘give away’ the meaning of the silence over the telephone, and thus one can assume silence to be performing positive functions, such as ‘good listening’, more readily. Steiner (1967, cited in Johannesen, 1974) offers a useful distinction, that silence is primarily non-verbalisation rather than non-symbolisation. Braithwaite (1990) and Samarin (1965, p. 115) argue that silence is indeed a symbolic resource in itself, not simply the absence of action or behaviour. As I shall show over the course of this thesis, not only is silence potentially meaningful, but it is also potentially not meaningful (i.e. inconsequential or not indicative of negativity). Despite the possibility of meaningful, responsive behaviour occurring during silence, these actions and the silence containing them have largely been outside the remit of traditional linguistic approaches. Saville-Troike (1985) calls for an integrated theory of communication that addresses the performative and communicative properties of silence, gesture, and language. This thesis will explore aspects of both of these issues and contribute to the development of one such theory using a conversation analytic approach.

### *[2.1] Contextual Considerations*

As Gardner et al. (2009) discuss, silence can vary across cultures, but this does not necessarily mean that cultures value silence differently in the grand scheme of things. It is possible, rather, that they have different orientations to interactions that are manifested in different degrees of silence. Stivers et al. (2009) looked at silences between yes/no interrogatives and responses; the ten languages that they studied show similar distributions for most of the languages but with significant differences in the means. Stivers et al. propose that speakers of different languages have different calibrations of silence. Allwood



(n.d., cited in Lehtonen & Sajavaara, 1985) hypothesises that cultures that have less tolerance for overlap have greater tolerance for silence and that cultures that have more tolerance for overlap have less tolerance for silence. This might be further dependent on sequential position, but preliminary evidence suggests that this is a reasonable hypothesis (Lehtonen & Sajavaara, 1985; Tannen, 1985).

However, this might not always be the case when further constraints on response timing are in effect. Nevile (2007) found that air crew members for major international airlines and air traffic controllers tended not to have silences at the boundaries of turns nor overlapping speech at any point in their interactions. Nevile (2006) also found that when precision timing does not occur between air crew and air traffic controllers, it is a possible sign that communication is impaired and that they may not be working together safely. In other words, ‘common but brief’ overlap (Sacks et al., 1974) that is observed in ordinary conversation was not only not observed but was extremely problematic for air crew and air traffic controllers. Although overlap was particularly indicative of problematic interaction, silence can lead to overlap. Silence can occasion participants beginning turns-at-talk, whether the turn is the sequentially implicated responsive action, a new initiating action, a repeat of the prior action, or pursuit of a preferred response. With all of these options available, participants sometimes begin speaking simultaneously following silence. Participants may begin speaking simultaneously due to possible problems with the preceding utterance<sup>4</sup>, for example, or because once silence sets in there is a lack of clarity when and who will (or ought to) speak next. In ordinary conversation and most institutions, however, silence and overlap do not tend to have such extreme consequences that interactional difficulties in aviation can have.

---

<sup>4</sup> This is true at least in minority world conversations. Philips (1976, p. 88) notes that people of the Warm Springs Nation rarely begin speaking simultaneously despite frequent long silences.

Although psychotherapy is a type of interaction that is *structured through talk* (Philips, 1985), psychotherapy is a situation in which one might expect considerable silences as a matter of course, rather than as a sign of trouble. However, little empirical research has looked at silence in psychotherapeutic or counselling contexts. Most publications dealing with silence in psychotherapy are written from prescriptive, theoretical positions by practitioners who are writing from their own interpretations of their own practice (e.g., Prince, 1997). Many of these are written from a psychoanalytic perspective in which the therapist is taught to allow silence to occur despite interactional relevance of therapist speech (e.g., Prince, 1997). In psychoanalysis, silence is taken to be meaningful and attributable to the client except when the psychoanalyst deems the client ready to receive an interpretation (Haim, 1990). Attributions of the meaning of the silence can purportedly be made from how the psychoanalyst feels about the silence her- or himself (Blos, 1972). Fliess (1949) argues that silence can be viewed as a form of sphincter closure to cause retention (of words, substituting for excretory products) and to punish the psychoanalyst in an act of transference of feelings about a dominant, cold parent to the psychoanalyst. In 1961, the idea that silence was a sign of aggression on the part of the client had remained salient to psychoanalysis. The *Journal of the American Psychoanalytic Association* published a special issue in this year on silence as a form of acting out (cited in Ephratt, 2008).

In other therapeutic approaches represented by writings on silence (e.g., crisis intervention, Scott & Lester, 1998; cf. Prince, 1997 from a psychoanalytic perspective), practitioners also express a struggle between silence and talk. Although authors that belong to other therapeutic camps do not tend to view silence as ‘urethral-erotic’, ‘anal-erotic’, or ‘oral-erotic’, as Fliess (1949) does, ‘resistance’ is a common theme, as is therapists’ discomfort with silence versus allowing opportunities for clients to speak. The conflicts

expressed by these therapists over silence in their practice may not reflect how naturally occurring silence in psychotherapeutic contexts is treated by clients. In fact, silence can be a resource for communicating certain elements of an issue that are not easily verbalised (e.g., that a topic is sensitive) or for which verbalisation would call into question the veracity of the statement itself (e.g., that one is sincere).

Silverman and Peräkylä (1990; see also Silverman, 1996, pp. 63-88) found that both patients and HIV counsellors used silence as one way to mark topics as sensitive. They describe this as ‘hesitation’, which includes perturbations such as ‘um’. The key point for them is that the substantive talk is being delayed. In other words, participants were actively but momentarily avoiding discussing or mentioning topics that were delicate, often involving sexual partners, sexual activities, or health statuses. In the psychotherapy data in my data set, disclosing or discussing information that is potentially sensitive to this degree is often not an issue in that the psychotherapist and client are well acquainted and the problems being discussed are of a less taboo nature. Chapter 4 of this thesis explores how silence can contribute to the performance of sincerity. I examine a particular sequential environment, clients responding to the therapist’s initiating actions. In this sequential environment, silences are used to demonstrate clients’ thoughtfulness or sincerity. Although the analysis of specific psychotherapeutic approaches is beyond the scope of this thesis, the therapist’s use of silences in formulating her own responses appears to be relevant to achieving a contemplative demeanour in the service of her spiritually oriented humanistic approach.

## *[2.2] Conversation Analysis and Silence*

Stivers & Rossano (2010a) critique the treatment of talk-focused interactions as a point of reference, noting that most of day to day human interaction does not involve long stretches

of talk. Rather, talk is something that occurs when interactants need or want to convey particular information. This information may be of practical or theoretical importance to the participants, or it may be ‘small talk’ for the sole purpose of social bonding. Silence outside of talk-based interactions is inherently unproblematic until attempts to engage in talk-based interaction fail. Silence within talk-based interactions, on the other hand, has been a subject of study per se within social research. Kincaid (1979) points out on the other hand that epistemological biases toward discrete messages (i.e. talk) has hindered our understanding of the role of silence in social interaction. In some contexts, silence may even be more acceptable than talk (Ishii and Bruneau, 1988; Murphy, 1970; Johannesen, 1974) in terms of an overall approach to interaction, a fact that is largely ignored by minority world scholars except to the extent that ‘other cultures’ can be compared to mainstream American and Western European cultures. The meaning for participants of a particular silence is evident in how they treat it interactionally and the sequential environment in which it is situated. Alternatively, silence can have no particular meaning in the interactional context in which it occurs (See Chapter 3). We must be careful to analyse silence according to how it is used and treated by interactants themselves.

Silence has been an issue for CA since the early years of its development with regard to turn-taking and preference organisation. In Sacks et al.’s (1974) paper on turn-taking for conversation, different functions and meanings of silences are discussed in depth. The sequential position of a particular silence can tell interlocutors and analysts whether a response, for example, is relevantly absent and which party or parties ‘should’ be speaking. A silence following an initiating action also signals that a dispreferred response may be forthcoming. In order to obtain a preferred response after a silence has begun, the speaker of the initiating action may rephrase the initiating action, produce a slightly different action (e.g., suggesting seeing a film instead of having dinner), or produce an alternative action

designed such that the same response is preferred according to the alternative action's format (e.g., suggesting that a person might have class that night). Note that 'dispreferred' refers to actions that are interactionally sensitive, for example a refusal, but not necessarily what the speaker of the initiating action wants or does not want. A person can invite another to a party but not really want them to accept (and the invitee can know that it is an insincere invitation). However, a quick, unmitigated refusal would appear ungrateful and abrupt despite being what the inviter actually wants. By constructing the refusal as regretful, both parties are able to maintain politeness. Chapter 2 of this thesis deals with methodological aspects of CA in depth.

With regard to the timing of silences, Jefferson (1988) suggests one second as an approximate 'standard maximum' for silence in conversation. However, as Gardner et al. (2009) and Stivers et al. (2009) demonstrate, one second is not a universal maximum, even in English. Chapter 3 deals with this issue using illustrative data from an American Mid-western dyad and supports Gardner et al.'s findings that the lengths of silence vary with situational contexts, even for the same participants. Chapter 3 additionally discusses implications for transcription accuracy in conveying meaning and technical considerations to readers who may or may not have access to supplementary recordings. Schegloff and Sacks (1973) propose that silences or lapses in interaction (which may be far longer than one second) in which closing and re-starting the interaction is unnecessary constitute a particular state that they term a 'continuing state of incipient talk'. Chapter 5 addresses this proposed state of interaction using content analysis of publications from 1973 to 2011 and critically reviews the usage and definition of this and related forms, such as simply 'incipient talk'. The content analysis identifies three distinct clusters of usage and definition, indicating that application of the construct is inconsistent and does not necessarily involve silences per se at all.

### **[3] Gesture Research**

Non-vocal behaviour can occur with or without accompanying talk, that is, without or with silence, respectively, and meanings may vary according to whether it is done silently. This thesis will focus primarily on non-vocal behaviour that occurs without accompanying talk. Research involving non-vocal behaviour without accompanying talk has focused on gestures (rather than, say, accomplishing tasks). The meanings of gestures can vary across and within cultures (i.e. across sub-cultures), and their placements and use by interactants are diverse and multi-faceted. Even so-called universal (Ekman & Keltner, 1997) gestures such as smiling can be performed very differently in different cultures (Pike, 1982, p. 51), including in different cultures of non-human primates (Preuschoft & van Hooff, 1997). Although humans may have the same vocal and non-vocal resources available, how they are organised and utilised in different languages and cultures varies (Sapir, 1925; Hall, 1969; Birdwhistell, 1968). Because gestures, like languages, comprise such a wide and multi-faceted area of enquiry, it is impossible to provide an exhaustive account of gestures and gesture research in the limited space available, even for English-speaking countries alone. In this section I will give a brief history of gesture research, not limited to stand-alone gestures but without an exhaustive account of gesture in human and non-human interaction. This section aims to orient the reader not only to current directions in interactional gesture research but that which underpins it as well.

#### *[3.1] Early Gesture Research*

Early gesture research focused on their use in rhetoric with a focus on prescription for successful oratory (e.g., Quintilian's *Institutio Oratoria*) in relation to thematic aspects of gestures' relationships to the ongoing talk. For many centuries, this approach dominated writings on gesture, and Bary (1679, cited in Kendon, 2004) produced a vocabulary of arm

and hand gestures and body postures based on prescribed interactional meanings of the gestures rather than relationships with talk (e.g., frankness, tenderness, complaining). This approach continued into the 18<sup>th</sup> century, although it is in sharp contrast with current approaches to gesture. Engel (1785) returned in some ways to earlier systems by beginning his system with a distinction between body movements that originate from bodily states and those that are related to emotion and purposeful activity. The latter category is further divided into categories based on function, which are each also divided into either precursors to intentional activities or figurative gestures.

Austin (1806) based his classification system on Engel's and deals extensively with hand gestures. Like classical writers on gesture, his primary focus was on gestures for orators. He looks at gesture from four viewpoints: the instrument (e.g., hand, head), signification (i.e. meaning), quality (e.g., boldness, simplicity), and style of delivery (i.e. the overall style including talk and its effects on the qualities of gestures). Most relevant to our analysis is the viewpoint of significance, which is divided into significant and non-significant gestures. Non-significant gestures are those that 'do not mark any particular sentiment; but are rather used to denote a sort of general relation in the expressions, and derive their significance from the time and manner of their application, from the place in which they are used, and from their various combinations', whereas significant gestures are those that are substantively communicative. Significant and non-significant gestures are then further subdivided similarly to Engel on a functional basis.

### *[3.2] Interactional Research*

Unlike previous writers on gesture, Wundt (1921, cited in Kendon, 2004) deals exclusively with gestures that can occur without accompanying talk. His system of hand gestures describes such gestures in terms of their symbolic function and their physical

implementation (or form). Wundt's move away from oratory marks a beginning in the movement toward an interactional focus in gesture research. Although some authors today have specifically focused on the semiotics of gesture (e.g., Morris, 1994; Calbris, 1990), the majority of gesture research today is experimental or interactional in nature.

Efron's (1941/1972) study of Eastern European Jewish and Southern Italian communities in New York City was perhaps the first truly interactional study of gesture and remains a key influential study in the field. Efron's aim was to systemically examine the gestures of first generation immigrants and assimilated descendants of both groups. He noted similarities between the gestures of the assimilated descendants and showed that gestures are cultural rather than genetic. This finding refuted nazi propaganda at the time that claimed that gestural styles were racially inherited. Although Efron's focus was on the similarities and differences of gestures across and within cultures, he deals with gesture from three viewpoints that contribute to a typology of gesture: spatio-temporal, interlocutional, and linguistic. The most widely known of these perspectives is the linguistic, within which famous categories such as 'emblem' are described.

Ekman and Friesen (1969b) offer a categorical system of gestures, some of which are based on Efron's. Their paper represents one of the formative works in 'nonverbal behavior' and 'nonverbal communication', and their work increasingly became concerned with detecting deception (e.g., 1969a). These pieces of research were of special interest to psychiatrists, and Ekman and Friesen's (1969b) system received a wide and supportive audience due to its well-articulated structure during the early days of this area of enquiry (Kendon, 2004, p. 95). Their system accounts for a broad range of body movements and is arranged around three perspectives: usage, origin, and coding. Gestures may belong to multiple categories within each area. For example, they provide a Venn diagram (ibid. p.



57) demonstrating how interactive, idiosyncratic, communicative, and informative categories can apply to the same behaviours. Ekman and Friesen recognised that behaviour may be intentional or unintentional and may have shared or individual meanings; they also recognised ways in which in the transmission of that meaning may function interactionally. They present five categories of 'nonverbal behavior' (emblems, illustrators, affect displays, regulators, and adaptors) as a work in progress that is not intended to be an exhaustive list nor exclusive categories but to build on the work of Efron with a view toward evolving specification.

Another influential perspective on gesture research is that of Ray L. Birdwhistell. Birdwhistell is famous for his framework of kinesics, which began with the application of a structural linguistics system onto non-vocal behaviour. The system he proposed borrowed terminology from structural linguistics to create concepts such as the kine and kineme (after concepts such as phone and phoneme). Unfortunately, because Birdwhistell was most concerned with teaching and avoided published anything he did not consider perfect, little work has been done in the field of kinesics (Kendon & Sigman, 1996). His ideas, have, however, stimulated the field of gesture studies. Adopting a structural linguistic framework in the development of kinesics was not a prescriptive approach but simply a starting point. Birdwhistell was concerned with the whole of co-present social interaction and saw it as a patterned system that could be analysed. This approach later became known as 'context analysis' (Scheflen, 1963, 1973).

Compared to more category-rich systems of gesture research, context analysis looks at the whole of an interaction, of which non-vocal behaviour is one part. Scheflen is one of the few who have employed kinesics as a methodology, and it is thoroughly incorporated into Scheflen's version of context analysis. Context analysis has been described by Kendon

and Sigman (1996) as a methodology ‘in which patterns of behavior in interaction are described and interpreted in terms of the contexts in which they occur’ and which are ‘aimed at systematic description of the total organization of behavior in interaction’ as opposed to those that study components of interaction in isolation from their situated use.

This definition of context analysis is more in line with Birdwhistell’s theoretical approach to interaction than how kinesics was and continues to be taken up by popular culture, academics, and practitioners alike – as ‘body language’. Although Birdwhistell approached non-vocal behaviour from a structural linguistics framework, he never intended that it would be considered language but rather a part of a cohesive whole that is human interaction (Kendon & Sigman, 1996, p. 245). ‘Body language’ is a misnomer that implies that non-vocal behaviours have specific meanings and are produced in syntactically appropriate ways to communicate consciously. Although the structural linguistics-based framework of kinesics is such that non-vocal behaviours may be combined to form *complex kinemorphs* that are equivalent to words or ideas, non-vocal behaviours are not placed such that a grammar or syntax per se exists outside of sign languages. However, non-vocal behaviours can stand in syntactic places that are structured by talk. Pike (1954/1967) describes a common party game at the time in which participants gradually replace words with signs that visually represent the words being replaced. He uses this example to illustrate how inattention to non-vocal behaviour (and in this case, specifically *non-vocal practices*, a term I use to describe those behaviours that perform specific interactional functions but which may or may not be gestures) can lead to misinterpretation of spoken language.

Pike’s major work *Language in Relation to a Unified Theory of Human Behavior* was written in three volumes and published between 1954 and 1960 with an updated

second edition in 1967. It covers a range of linguistic issues including non-vocal behaviour. Although Pike primarily studied the languages of indigenous societies (drawing heavily on his teacher Edward Sapir), and devotes much of volumes two and three to this endeavour, the first volume of this work covers more general linguistic issues. Volume one deals with most of the issues that are directly relevant to non-vocal behaviour as an area of study. It is also in this volume that he coins the terms *etic* and *emic*. Etic refers to perspectives or analyses that are intended to treat all groups equally and is characterised as generated from an outsider's perspective. It is sometimes described as objective, as Pike himself on occasion does in his description of it. However, Pike's use of 'objective' does not imply a realist or empiricist objectivity. Rather, Pike's notion of objectivity is simply that of an outsider's judgement along with all of the values and experience that that outsider might bring. Multiple researchers may describe the same phenomena differently yet all come from an etic position. Emic, on the other hand, describes perspectives or analyses that reflect what is meaningful to participants, rather than based on the outsider's perspective. Although an outsider may describe a phenomenon, they may do so from an emic perspective by finding out what is meaningful to those involved. Etic and emic do not represent discrete categories of descriptions but rather are two extremes on a continuum. Pike states that when a person first approaches a group with whom they have not had contact before, they will always have an etic perspective. As they become more familiar with the values and practices of the group, they will develop an emic perspective.

Pike's concepts of etic and emic are useful for the understanding of conversation analytic approaches. A key feature of CA is that it is able to describe both structural regularity across cases (i.e. an etic perspective) as well as how individual parties in particular interactions orient to structures and interactional resources (i.e. an emic perspective). CA strives to marry these two perspectives to achieve a multi-level

description of social interaction (Schegloff, 1987). Although CA's methodology is based primarily in talk, the flexibility of CA in examining particular interactions and corpora has relevance for the study of non-vocal behaviour. As I have mentioned above, the meanings and use of non-vocal behaviour are dependent on a variety of factors such as the immediate sequential environment and the broader cultures to which the interactants and setting belong. Both issues must be considered in the interpretation of individual instances. However, the broader structural features of interaction are often quite regular across participants and situations. In discussing individual cases, analysts are able to demonstrate how a practice can be used by interactants as well as some of the effects that conversational organisations have on interactants' actions (Schegloff, 1987).

Whilst many interactionally based researchers take the perspective that interaction must be studied as a whole (e.g., Pike, 1967), some take a more language-focused approach to gesture. McNeill (1985, p. 350) makes the (perhaps still) cutting edge statement that the designation of 'linguistic' as what we can write down and non-linguistic as everything else 'is a cultural artefact and an arbitrary limitation derived from a particular historical evolution.' This is in itself a bold statement against ethnocentrism and ableism for his time. McNeill, in proposing a new way to view the study of gesture, asks us to question conventional minority world linguistics and psychology of language assumptions. He goes on to provide an account of why, in his view, gesture is *verbal*. Whilst we<sup>5</sup> disagree with McNeill on many of his arguments supporting this claim, he presents a revolutionary and radical approach for the time in which *So You Think Gestures Are Nonverbal?* was written.

---

<sup>5</sup> Although this thesis is my own work, I use 'I' and 'we' throughout the thesis. Portions have been published or are under review as co-authored papers, and I have retained 'we' to be true to the environment within which the research was produced.

He argues that gestures, including 'referential' and 'discourse-oriented', are 'verbal', because '(a) Gestures occur only during speech, (b) they have semantic and pragmatic functions that parallel those of speech, (c) they are synchronized with linguistic units in speech, (d) they dissolve together with speech in aphasia, and (e) they develop together with speech in children' (ibid. p. 353). These claims are, however, weakened by a number of pieces of evidence:

(a) He claims that gestures *only* occur during speech production (ibid. p. 353), despite also claiming that 'more than 90%' occur during speech production (i.e. that approximately 10% of gestures do not occur during speech production) (ibid. p. 354).

(b & c) Whilst he makes a strong case that certain gestural patterns are associated with and usually synchronised with speech production (ibid. p. 354-361), he provides no evidence that they are necessarily or inherently intertwined (but see later work on the link between speech production and spontaneous gesture, such as Iverson & Goldin-Meadow, 2001).

(d) He lumps aphasia into one broad category (ibid. p. 361-362), whilst clinical and experimental evidence shows strong differences between types of aphasia and their effects on gesture and speech (See Feyereisen, 1987 for a specific critique of McNeill's treatment of aphasia). Moreover, recent cognitive neurolinguistics research has shown that people do not have uniform neural pathways for language and information (Łojek, 2009). The elaborate use of gesture by people with aphasia (e.g., Goodwin, 2003, 2004, 2006, 2011) also suggests that at least referential gestures may in fact increase with aphasia, as people try to find other ways to communicate.

(e) While he reviews the literature on gesture and language development in children more thoroughly than aphasia, there is no clear indication that gesture and language develop synchronically in children, but rather that physical play may be a precursor to gesture and not necessarily to language (ibid. p. 362-365). The youngest gesturer they recorded was 5 years old, far older than a child who can be expected to use language. See Butterworth &

Hadar (1989) for a thorough critique of McNeill's assumptions regarding language development.

One of the most common aspects of modern gesture research is the recognition that communication is not merely a product of multiple parallel channels providing different information. Rather, information is communicated as a combination of different kinds of resources and modalities that moderate one another. This is particularly evident in the work of Birdwhistell and Pike, but prominent gesture researchers such as Kendon (1994) continue to address this issue. Kendon takes a broad, social psychological approach to gesture research and displays a tendency to examine particular gestures or classes of gestures in how they are used in social interaction (e.g., Kendon, 2002). Kendon's interactional focus in his empirical work parallels that of conversation analysis in some ways, such as the use of detailed transcripts to illustrate phenomena and a sequential orientation to some problems. Kendon's work is variably concerned with conversational structures (e.g., Kendon, 1997, p. 114), interactional frameworks (e.g., Kendon, 1988a), and semiotics of gesture (e.g., Kendon, 2002).

### *[3.3] Conversation Analysis and Gesture*

CA began through the study of telephone calls and audio-recorded co-present interactions. Focus on talk in the early days of CA helped to make the actions performed by parties explicit. However, both Goffman and Garfinkel, who influenced the development of CA in substantive ways, were concerned with broader issues such as how people engage with one another (e.g., Goffman, 1967) and how social institutions are constituted by the people within and interacting with them (e.g., Garfinkel, 1967). Nevertheless, most research within a CA framework on non-vocal practices has focused on gestures that occur with

accompanying talk. Whether the (continued) focus on talk developed from engagement with the opportunities afforded by the data available at the time or whether it started with a talk-focused ideology cannot be answered conclusively. However, the ideology does not have to remain. The focus on action formation that underpinned early sequence organisational studies is a useful way to approach gestures and other non-vocal practices. By having an existing body of research on how talk is structured, the task of studying non-vocal practices becomes simplified. Because non-vocal practices do not usually constitute the whole of an interaction, their functions in alternation with and alongside talk become visible through their sequential positions, originally identified through work on talk.

The research of Jürgen Streeck and Charles and Marjorie Goodwin has been particularly influential in conversation analytic work on non-vocal behaviour. Streeck's approach is cross-cultural and deals primarily with speakers' hand gestures in co-present interactions with adults who do not have communication difficulties (e.g., Streeck, 2009a). Charles Goodwin's work on non-vocal behaviour concentrates on the communicative resources available to people with aphasia (e.g., Goodwin, 2004), and Marjorie Goodwin's work has focused on the communicative resources of children (e.g., Goodwin, Goodwin, & Yaeger-Dror, 2002). Research on gestures and other non-vocal practices from a CA perspective has, as mentioned previously, concentrated on how non-vocal practices facilitate or enhance talk.

Lerner, Zimmerman, and Kidwell (2011) have, conversely, identified task structures as an area of sequentiality to which even very young children orient. Lerner et al.'s analysis is built primarily around the example of a mother and small children preparing for lunch. They demonstrate how the participants orient to phases of the task and sequentiality of actions within the task in designing turns-at-talk and performing non-vocal behaviours.

In their words, the ‘episodic and formal phase structure’ is a ‘resource for action’. Lerner et al.’s findings both enhance and challenge conversation analytic thought on non-vocal practices. Non-vocal practices (particularly tasks) can themselves be sequentially structured resources. In Lerner et al.’s data, non-vocal practices function somewhat external to the organisation of talk, but the process of preparing for a meal is an interaction that is *structured through silence* (Philips, 1985). Although their analysis lends support for viewing non-vocal practices as sequential resources, they do not definitively address tasks within what conversation analysts commonly know as *sequences*, that is, actions built around pairs of actions and expansions upon them.

Sacks and Schegloff (2002; cf. McNeill, 1992, p. 83) describe the internal phases of gestures that occur concurrently with talk and show how gestures reach an apex before returning to ‘home position’. The interactional relevance of this structure has been further discussed by Cibulka (2012), who shows that gestures can have multiple home positions at points along the way and that these intermediate home positions function differently from the initial home position. The intermediate home positions maintain engagement and readiness to respond, whilst the returning to the initial home position indicates completeness or withdrawal. As with Lerner et al.’s (2011) work, Sacks and Schegloff (2002) and Cibulka (2012) address structural concerns that do not map directly onto the sequential structure of conversations but which indicate that *some* level of sequentiality is at play.

Individual non-vocal resources have been described in terms of their sequential positions within interactions that are structured through talk. For example, gestures that are performed by one party and are then repeated by another party during a subsequent turn-at-talk (‘return gestures’, de Fornel, 1992; Koschmann and LeBaron, 2002, p. 262) are used



to display intersubjectivity. Interrelated vocal and non-vocal practices can be used by participants in alternation with each other within the same turn-at-talk as well. Olsher (2004) describes how actions begun by talk can be completed by gestures. Little research has examined the functions of non-vocal practices that constitute entire actions, however Seo and Koshik (2010) and Mortensen (2012) have identified a small number of gestures that function as repair/correction initiators and regularly receive verbal responses in foreign language classrooms. Additionally, Whitehead (2011) has described variations on nodding that are used by the speaker of an initiating action following a response. Although specific non-vocal resources with apparent sequential positions and that have implications for the ongoing interaction have been identified, the broader role of non-vocal practices within sequence organisation remains unexamined.

#### **[4] Conclusion**

Although silence is often perceived in minority world cultures as inaction or inactivity, people may perform many communicative activities during or through silence. These activities can emphasise ‘doing nothing’, or although the interactant is not progressing the interaction verbally, something else is relevantly occurring (e.g., ‘doing thinking’ or non-vocal reactions). If they are in fact doing *something*, to quote Sacks and Schegloff (1973), ‘why that now?’ Although the structural organisation of non-vocal practices has been addressed in terms of socially relevant internal sequences, the sequential role of non-vocal practices within interactions that are structured through talk has only been examined in terms of particular practices. How, then, do non-vocal practices function within sequences? In order to understand the role of non-vocal practices such as gesture, we must first understand how silence, which has traditionally been conceptualised as occupying sequential space as the absence of talk and therefore absence of sequential action, is treated by participants. In other words, how does silence function within the sequential

environment? How does context affect silence and non-vocal practices? If one organisation (or aspect of an organisation) is affected, are all organisations (or aspects of that organisation) affected?

In the chapters that follow, I will explore how a comprehensive CA approach to silence and non-vocal practices in co-present interaction can clarify the status of particular cases as inaction or silence action. I will show how practices that have been methodologically treated as inactivity in fact perform very specific and delicate functions. The focus of my thesis will be on sequence organisation and, to some extent, turn-taking. Chapter 2 will examine CA in further detail and establish its appropriateness as a methodology for the study of social interaction, including silence and gesture and provide an introduction to the method. In Chapter 3, I will examine how length of silence can vary in interactions involving the same participants and discuss the appropriate analysis and transcription of lengthy silences across situational contexts. Building on this contextual theme, in Chapter 4, I will present one context in which silence operates in a way that is different from that described in ordinary conversation, namely how silence can contribute to the performance of sincerity in psychotherapy. Unproblematic silence is sometimes accounted for with the concept of ‘continuing states of incipient talk’, which is a term with varying definitions. Chapter 5 will address the problematic nature of this term, including divergent uses and lack of empirical evidence. This chapter will include a cluster analysis of how incipient talk has been used and defined as well as possibly related concepts and benefits and drawbacks of those concepts. One usage of ‘incipient talk’ involves non-applicability of sequence organisation, and a context that is commonly considered a continuing state of incipient talk is that of television-watching. Chapter 6 will involve television-watching and its effects on sequence organisation. I will show how television-watching involves a relaxation of certain aspects of sequence organisation but strong

orientation to turn-taking and certain other aspects of sequence organisation. Television-watching is one environment in which non-vocal practices might be expected as a matter of course to avoid talking over the television, but interactants may produce non-vocal practices in contexts in which sound competition is not an issue. I will describe some uses of gestures in responsive actions in Chapter 7, and in Chapter 8, I will argue for the inclusion of non-vocal practices in sequence organisation in CA methodology. Finally, in Chapter 9, I will discuss implications of the research presented and make recommendations for future research.

## **Chapter 2**

### **An Overview of Conversation Analysis and the Data Presented**

**Abstract**

In this chapter, I will describe conversation analysis (CA) methodology and establish its suitability as a method for studying silence and gesture and provide an introduction to the method including history and applications. I describe the data from which this thesis draws. I then discuss key methodological issues in CA including cultural intelligence, accurately conveying information, and organisational structures around which CA is built: overall structural organisation, sequence organisation, action formation, turn taking, preference, trouble/repair, world selection, and topic. I then discuss CA in relation to silence and non-vocal practices and argue for why CA is well-suited to the aims of the thesis.

**[1] Introduction**

Conversation analysis (CA) is a rigorous qualitative method that examines the ways in which interaction progresses and how that progression might be affected by participants' actions. Meaningful social research should both specify (describe accurately) and account for (give reasons for) the phenomena under study. CA both specifies and accounts for the structures, practices, and actions that make up our real time interactions with other people, whether we are co-present or interacting via a telecommunication system and whether we are speaking/signing directly to one another or communicating with an aid such as a notepad or an interpreter. So far, CA has focused on mainly vocal (and specifically verbal) elements of spoken interaction, but with the widespread availability of video recording equipment, non-vocal practices as well as other forms of non-vocal communication (such as the use of augmentative and alternative communication [AAC] devices) have been given more analytic attention (e.g., Bloch & Wilkinson, 2004; Korkiakangas, 2011). Topics of interest to conversation analysts thus far have ranged from structures with far reaching influence (e.g., turn taking, Sacks et al., 1974, and repair, Schegloff, Jefferson & Sacks, 1977) to practices with more specialised applications (e.g., antibiotics prescribing, Stivers, 2005a,b). This chapter aims to give a brief account of the development and applications of CA, describe the data used in this thesis, consider the application of CA to psychotherapy (from which a large portion of the data has been obtained), describe key methodological constructs and systems, and consider the application of CA to silence and non-vocal practices.

**[2] History of Conversation Analysis**

Conversation analysis has evolved rather quickly in the past half-century from a basic (rather than applied), socio-linguistic method to study interaction (through such topics as the nature of greetings and closings) to an interdisciplinary methodology with complex

transcription conventions. Conversation analysis emerged from the early work of Harvey Sacks and Emanuel Schegloff. Sacks (1995, p. 3) was asked by the mental health centre with which he was working to assist them in developing ways of getting names from callers to their suicide prevention hotline. In effect, conversation analysis's roots are in applied research despite early publications of a more basic nature. Schegloff similarly worked with telephone calls to the police at this time, which resulted primarily in his classic, basic work on telephone call openings (Schegloff, 1968).

CA is a data-driven technique that relies on recordings of naturally occurring interaction. Around the time of the inception of conversation analysis, there were other sociologists who worked from the 'bottom-up' as well. Some of this work – that of Erving Goffman and Harold Garfinkel – formed the foundation on which CA was built. Goffman advocated for social order as an object of study with a focus on the workings of social interactions. Primarily theoretical, his work explores a number of issues related to how people come to interact, how they manage their participation in interactions, and how public and private engagement can co-exist and change. Garfinkel, on the other hand, approached the social lives of humans with a wide lens, looking at how people construct and perform their roles, institutions, and cultures. Although Garfinkel looked at how people accomplish what they set out to accomplish socially, his research was less concerned with the workings of conversation per se and interactional frameworks than Goffman's. Garfinkel's development of ethnomethodology was in reaction to the researcher-led nature of early sociological research, and he was one of the first authors to advocate as a programme of research that participant orientation is essential to understanding the social order. He imparted to CA the value of data-driven basic research that is reported (ideally) without political or psychological motivation.<sup>6</sup>

---

<sup>6</sup> I am grateful to Y. Gavriel Ansara for discussions of ideologies behind Garfinkel's (1967, p. 116-185 and p. 285-288) accounts of Agnes.

CA was developed through deep consideration of the nature of sociality that was influenced by the thinking of both Goffman and Garfinkel, as is particularly evident in the Sacks (1992) lectures. This consideration led Sacks and Schegloff (and later Gail Jefferson) to not only incorporate aspects of Goffman's and Garfinkel's respective approaches but to rely on recordings of naturally occurring interactions, which maximises the analyst's ability to accurately transcribe data and address participant orientation by not relying on memory. CA aims to describe the process by which real-time social interaction unfolds (Psathas, 1995, p. 7). This is accomplished primarily through turn-by-turn analysis, meaning that each action (usually a turn-at-talk) in interaction creates openings for, and limits, possible responses. Thus individuals build the meaning of an interaction for themselves, not only at the outset but as the interaction unfolds.

Some researchers who also worked with naturally occurring data around this time, such as Robert Bales (e.g., 1950; 1969) and George C. Homans (e.g., 1958), utilized theories of social behavior and coding techniques based on theoretical or pre-existing constructs. CA, however, has retained its bottom-up approach throughout its history. Additionally, basic conversation analytic research, or 'pure CA', tends to dispute the analytic relevance of information recalled by participants such as thoughts before, during, or after an interaction as potentially informed by the interaction or assumed from participants' expectations of what they ought to think, do, or say. Pure CA tends to avoid assuming the relevance of institutions and social roles, but rather to let such issues emerge as interactionally relevant from the participants themselves. This is not to say that in order for social roles or institutions to be relevant, people must explicitly state or reference a role or institution but that they may orient to them (or not). CA avoids a priori assumptions about relevance, but the approach does not preclude relevance.



Sacks's data from the Suicide Prevention Center were central to his first lectures at UCLA (compiled in Sacks, 1992) and are the first conversation analytic transcripts despite having been written in standard orthography. A later colleague, Gail Jefferson, was the primary force in transforming conversation analytic transcription (e.g., Jefferson, 1983). This transcription style is now termed *Jeffersonian* transcription. As the purview of CA has broadened from telephone calls and audio-only recordings of co-present interactions to non-vocal behaviour and 'paralinguistic' aspects such as prosody, so have the transcription needs of conversation analysts broadened. Transcription styles today are often based around Gail Jefferson's transcription style but are adapted specifically for the purpose of the research. For example, Goodwin (e.g., 2004) often provides photographs or drawings of relevant interactional details such as gestures within written transcripts.

### **[3] For what research questions is conversation analysis useful?**

CA has been useful in research on telephone calls (Sacks, et al., 1974; Schegloff & Sacks, 1973), flight decks (Nevile, 2004), medical consultations (Heritage & Maynard, 2006), psychotherapy (Peräkylä, Antaki, Vehviläinen, & Leudar, 2008), playgrounds (Goodwin, 1985), mealtimes (Goodwin 1979), among many other topics. How, however, would CA be any more appropriate than another method to address issues in these or any other setting?

Conversation analysis is uniquely poised to address research questions that:

#### *a) Are inductive*

Conversation analysis is much better suited to answer 'how' type questions such as 'How do people deal with problems of hearing?' than yes/no type questions such as 'Do students work harder in the presence of a teacher?' CA begins with the interaction rather than a set

of theoretically based codes or categories. Conversation analysts are certainly not any more objective as people than any other researchers. However, they attempt to let the data speak for itself. This is most noticeable at least during the first stages of analysis, but data extracts are essential to the dissemination of conversation analytic research. Data sessions (meetings during which conversation analysts look at a piece of data in great detail) are useful for identifying practices that can then be explored in a more methodical manner. This will lead to a main collection or collections and smaller collections of related instances. The analyst will then ideally be able to explicate the differences and similarities among these collections and give an analysis of how each practice works across the data. Generally, a practice is called a 'resource' in basic literature, and resources are utilised to build actions. An argument for the use of 'practice' rather than 'resource' is that some practices may have negative outcomes, which would make it quite odd to call them 'resources'. 'Practice' is also in line with the terminology of clinical fields, and some of the work presented in this thesis is aimed at those who teach communication skills to clinicians or are clinicians themselves. For these reasons, I have chosen to use primarily 'practice' rather than 'resource' despite the tendency to the contrary in much of CA literature.

*b) Are based on naturally occurring interactions*

Because conversation analysis looks at how people bring about and orient to structures in interaction, naturally occurring interactions are of primary concern<sup>7</sup>. Some interactions contain scripted components (such as interview questions as with telephone surveys or qualitative interviews or call scripts as with many customer service or telemarketing agencies). However, they are naturally occurring in that the interlocutors must work out for themselves when to speak next, whether the other has adequately answered (or asked) a question, etc. and deal with problems of hearing or understanding.

---

<sup>7</sup> Much work has been done on interviews of various types (e.g., Heritage & Roth 1992 on news interviews).

*c) Involve interactions that would not be dangerous to record for any party involved*<sup>8</sup>

Although it may be obvious that one would not want to have a camera in certain situations, for example in studying street gangs, this is an issue which one must consider when making the decision to do conversation analysis over another method. Ethnography is certainly a better medium for studying street gangs in their neighbourhoods, but conversation analytic work has been done on police-suspect interactions (e.g., Stokoe, 2009). There are concerns which must also be addressed in less overtly risky situation regarding moving people or objects, confined spaces, or moving vehicles, but in general any interaction which can be recorded in some safe manner can be data for conversation analytic research. For example, it might not be safe to hold a camera or microphone during a spelunking or mountain biking trip, but a miniature camera and/or microphone fitted to one's helmet or clothing could provide a viable alternative.

*d) Grounds analyses in participants' observable behaviours rather than internal processes*

Except as they are displayed in the interaction, attitudes, thoughts, and identities are all internal processes. How they are displayed or become relevant in interaction may be quite different from what one would tell a researcher or what one actually feels. How they are displayed and in what situations they become relevant may be of interest to a conversation analyst. For example, Land & Kitzinger (2005) found that heterosexuality was often assumed in telephone calls with businesses, such that it was relevant for lesbian callers to mention their non-male partners. Land & Kitzinger showed that did not tend to initiate repair which interrupts the progressivity<sup>9</sup> of the talk (e.g., 'And what is your husband's name?' 'Actually, I'm a lesbian. My wife's name is Lucy.') but rather made corrections which do not (e.g., 'And what is your husband's name?' 'Her name is Lucy.'). The researcher may opt to ask about internal processes, particularly concerning identities or

---

<sup>8</sup> Some environments can be risky in their own right but not specifically to record. For example, there is always risk involved in piloting an aeroplane. However, recording the interactions on the flight deck would not increase the risk and is regularly done in the interest of safety.

<sup>9</sup> Progressivity refers specifically to the process of completing courses of action, rather than the general notion of moving a conversation forward.

attitudes, but this is generally outside the scope of conversation analysis as a method.

Thought processes in particular are considered to be highly influenced by memory and post hoc intellectualising, so they would most likely not make it into such lines of questioning.

They may, however, make interesting field notes that may further the researcher's ideas on a topic.

#### **[4] Context of the Thesis**

My thesis was originally comprised of a study called *Good Relations*, which was focused on medical and psychotherapeutic interactions involving adults with self-designated genders (who might identify as or be considered by others to be 'trans' or 'intersex') and comparative data involving adults with assigned genders. Although data were collected and publications generated as part of the *Good Relations* project, the thesis focus has shifted to address broader, methodological issues in the analysis of silence and gesture. Certain challenges that were faced during the *Good Relations* project made it untenable as a time- and scope-limited PhD project, which I detail below to contextualise the inclusion of multiple data sources in this thesis.

Direct mailing and listservs were used for *Good Relations* recruitment of individuals and service providers. Recruitment was largely focused on medical and psychological care professionals with a large service proportion of adult service users with self-designated genders in order to obtain comparative data with service users who had assigned and self-designated genders. A number of challenges prevented full realisation of the *Good Relations* study within the resources and time-frame of the PhD. UK Psychotherapists and counsellors were generally unsupportive of the research, citing client confidentiality as their main objection and preventing clients from deciding for themselves whether to participate. This did not in itself constitute an insurmountable hurdle to the project.

Physicians and surgeons were generally very supportive of the research, as were potential patient and client participants. The National Health Service (NHS) ethical approval process was lengthy yet relatively simple. However, the NHS research and development approval process presented significant challenges to the study as one that was intended to recruit service users and healthcare providers *where they were* rather than from a particular area. Because the Research and Development Office for every NHS trust involved had to approve the research and individual researchers, and each office had different standards, the only way to recruit NHS participants would have been to secure participation commitments, apply to the particular Research and Development Office using their particular system, potentially apply for a Criminal Records Bureau check for each new office, and await approval – potentially to be told the research had to change from what had been approved by the Research Ethics Committee. It was thus impractical to recruit in a way that would be convenient to participants, particularly busy healthcare providers whose availability or ward arrangements could change on relatively short notice. Ultimately, the Research and Development process represented a substantial challenge that made data collection in UK settings unfeasible. I have thus only used a limited amount of *Good Relations* data that was collected in a US psychotherapeutic context, which was subject to University ethical review and participant informed consent (detailed below) rather than NHS Research & Development procedures.

### **[5] The Data**

The data that I have collected and used in this thesis come from everyday and psychotherapeutic interactions. The everyday data was collected as part of the *CA Data Bank* at the University of Roehampton under the supervision of the Principal Investigator Dr John Rae, whilst the psychotherapy data was collected as part of the *Good Relations* project described above. *CA Data Bank* participants were recruited by word of mouth, and

the project was open to all people 18 years of age and over. The *CA Data Bank* participants consented to the sharing of their data with researchers in general or specifically with me. In addition to the data that I have collected, *CA Data Bank* data that were previously collected by John Rae and by Terhi Korkiakangas have been utilised. Rowena Viney has also provided some of the data extracts that are analysed in Chapter 6 *How Television Plays into Sequence Organisation*. All of the data in my thesis comes from participants who consented to be video recorded, to have these recordings analysed, and to the publication of pseudonymised transcripts of their interactions.

The therapist and clients who were recruited under *Good Relations* were recruited by word of mouth. The therapist provided the clients with the recruitment materials, because she judged that these two clients might be interested in the project and would be unlikely to feel pressured to participate to please her. Both clients and the therapist expressed that they were enthusiastic about the research and its potential to help therapists improve their practice. They were also generally relaxed during data collection, as evinced by their general manner as well as joking about the camera. All queries and informed consent procedures were handled by me in order to limit any obligation clients might have felt to the therapist.

*Good Relations* and *CA Data Bank* participants were lent consumer-grade camcorders and were able to decide when and what to record, for how long, and whether to erase any portion of the tapes. They were informed that they could withdraw at any time without giving a reason. All data is stored in hard copy in locked cabinets or in electronic format on encrypted discs. The following sub-sections detail the data that has been used to illustrate analyses but does not include all data that has been collected as part of the *CA Data Bank* at the University of Roehampton. Appendix B contains the consent form for the

*CA Data Bank*, whilst Appendices C-G contain the ethical approval (Appendix C), recruitment letters (Appendix D), information sheet (Appendix E), consent form (Appendix F), and debriefing letter (Appendix G) from *Good Relations*. The data provided by Rowena Viney were collected as part of her PhD research on lesbian household interactions, which was originally approved by the University of York ethics committee (see Appendix H for her consent form and Appendix I for her written permission to use the extracts in this thesis) and which has since been continued at Loughborough University.

In individual chapters, I have focused on different aspects of these settings. In particular, I have primarily used data from interactions between a psychotherapist and individual clients *during* psychotherapy and familiars watching television or films at home. These are two interactional environments in which radically different orientations to talk might be expected, and which indeed exist in this corpus. They are therefore the most applicable in the corpus to address participant orientations to silence and gesture that differ from canonical conversation. In the sections below, I introduce the reader to the participants in the data that has been used in the thesis.

### [5.1] *Coral & Al*

The data involving Coral and Al comprises approximately 20 hours of video-recordings of household interactions including television-watching, cooking, meal times, and other household activities. Coral and Al live in the southern United States and have lived together as a couple for approximately five years. Coral is a middle-aged woman of mixed (visibly European) descent, whilst Al is a middle-aged man of mixed (visibly African) descent. Coral and Al participated in the *CA Data Bank*.

*[5.2] Paula & Leif*

This data comprises approximately six hours of video-recorded psychotherapy sessions between Paula and Leif. Paula is a spiritually oriented humanistic psychotherapist, and Leif is a client who has been seeing her for approximately three years for assistance with recovery from alcoholism. He also has mobility difficulties (which cause him significant anxiety) as well as difficulties with his wife. These are the main issues that are addressed in the therapy sessions. Leif is a middle-aged European American man, and Paula is a middle-aged European American woman. They are both from the western United States. Paula and Leif participated in Good Relations.

*[5.3] Paula & Sally*

The Paula who is featured in this data is the same therapist as is featured in the data with Leif. This data also consists of video-recorded dyadic psychotherapy sessions between a therapist and client. Sally is a middle-aged self-designated<sup>10</sup> woman who has Parkinson's Disease and is seeing Paula primarily for assistance with her gender affirmation and the difficulty her family has with accepting her. Paula and Sally participated in Good Relations.

*[5.4] Kim & Barbara, Dana & Niles*

The video data involving the dyads Kim & Barbara and Dana & Niles were collected by Rowena Viney and has been shared for the purpose of illustrating particular analyses. Both dyads are roommates in different households in England in which at least one member identifies as a lesbian. In the case of Dana and Niles, both are English, and in the case of Kim and Barbara, Barbara is English and Kim is American. In the extracts that I use, participants are watching television or a film at home, as Chapter 6 stemmed from a

---

<sup>10</sup> *Self-designated* is an empowering term to describe genders that people designate for themselves and may differ from those that they were assigned. Some people with self-designated genders may identify as or be considered by others as transsexual, transgender, genderqueer, or intersex, although this is not always the case. (Ansara & Hegarty, 2012)



collaborative project on social interaction during television-watching (Berger & Viney, 2010; Viney & Berger, 2010).

*[5.5] Kirsi & Elisabeth*

The data involving Kirsi and Elisabeth were collected by Terhi Korkiakangas (2012) as part of the CA Data Bank. They are used for the purposes of illustrating particular analyses, because they provide the clearest examples in the available data of particular, uncommon phenomena. Kirsi is a woman in her twenties, and Elisabeth is her grandmother who has been diagnosed with dementia and Parkinson's Disease (Korkiakangas, 2012). Both are Finnish, and the videos were recorded in Finland.

**[6] Application of Conversation Analysis to Psychotherapy and Related Fields**

Although this thesis does not focus on psychotherapy as an analytic issue overall, it is necessary to explore the analytic relevance that psychotherapy as a setting can present for CA. Perhaps one of the most innovative early conversation analytic healthcare studies was the 1990 Silverman and Peräkylä study *AIDS Counselling: The Interactional Organisation of Talk About 'Delicate' Issues* (Silverman & Peräkylä, 1990). Nearly all social scientific enquiries about HIV before and shortly after theirs had involved epidemiology or health knowledge and behavior through survey methods. Silverman and Peräkylä instead offered an account of how service users and healthcare professionals in HIV counselling settings organised their talk in relation to 'delicate' issues, in particular the function of delay in setting up the topic as 'delicate'. Silverman (1996) later went on to provide an entire volume on talk in HIV counselling settings. This HIV counselling research was unique not only in its time but for many years. Much of the conversation analytic healthcare research since Paul ten Have's 1987 book has been on GP consultations and the general notion of patient-provider relationship (e.g.,

Coupland, Robinson & Coupland, 1994; Heritage, 2005; Heritage & Robinson, 2006).

Psychotherapy is becoming an increasingly common subject of study within CA (e.g., Peräkylä, Antaki, Vehviläinen, and Leudar, 2008).

Although some of the first recordings used for CA were from a suicide prevention line and a group therapy session, the interactions were not analysed in their therapeutic capacities. HIV counselling certainly goes beyond mere recommendations for medication and touches areas such as sexual behaviour, stigma, and emotion, and Peräkylä & Silverman clearly had its therapeutic capacity in mind. However, this setting on the fence of medical and psychological care would not immediately give rise to what has become an important area of study in conversation analysis. Indeed, there is not a move to segregate medical and psychological care clearly in CA. There are many settings in which the two topics are inherently intertwined. For example, Kitzinger & Kitzinger (2007) look at a hotline for women who have experienced trauma from giving birth and may have ongoing medical and psychological needs and who may be currently in crisis. Surprisingly, there is a substantial lack of research in the area of psychiatry despite a growing body of CA psychotherapy research. Speer and Parsons (2006, 2007), however, recently published two papers on the interactions between psychiatrists and men and women of transsexual experience at an English 'gender identity clinic'. This is a highly specialized field that does not generally involve prescribing of psychiatric medication, but rather assessing people's readiness to affirm their genders, which in many places does not necessarily involve mental health professionals.

Conversation analysis in psychological care has only been an emerging field of the past decade. Madill, Widdicombe, and Barkham (2001) demonstrate the potential of conversation analysis with regard to psychotherapy, particularly in understanding how

matching of patients and providers affects the patient-provider relationship. Sutherland and Couture (2007) demonstrate that conversation analysis can be useful in understanding how alliances during family therapy are formed and affect the session. Peräkylä, Antaki, Vehviläinen, and Leudar (2008) published a collection of CA studies on psychotherapy. Whilst many of the studies in this volume are on psychoanalysis, it has some relevance for a wider audience and has contributed to greater interest in the area. For example, Leudar, Sharrock, Truckle, Colombino, Hayes, and Booth (2008) describe the sequential structure of psychoanalytic play therapy; the strategies that these therapists use to engage the children can be useful to other play therapy perspectives. Other studies in this volume, such as Rae's (2008) work on lexical substitution, describe ways in which clients and therapist take advantage of the conversational resources available to them to accomplish interactional tasks in the therapy session. Chapter 4 of this thesis takes a similar approach in looking at silence in the performance of sincerity.

## **[7] Doing Conversation Analysis**

### *[7.1] Cultural Intelligence in doing Conversation Analysis*

A certain degree of cultural intelligence is required of the analyst to make reasoned analyses of the dynamics of the interactions that they study (ten Have, 1990). This competence would normally be present in a person studying everyday interactions in a familiar language, region, and/or sub-culture. As most conversation analysts have backgrounds in the social, linguistic, or psychological sciences, rather than business, law, medical, or other field that they study, the cultural intelligence that a psychotherapist, for example, would have almost automatically needs to be learnt for the purpose of study. From its inception, CA has focused on the structures and practices of primarily verbal interaction. As it has become more applied, many conversation analysts have given more consideration of contextual factors, primarily so that the issues and terms raised in the

interaction can be more readily accessible. This learning can be accomplished by formally observing the setting where data will be collected, building on prior knowledge by asking questions of people in the particular field or setting or other potentially knowledgeable parties such as other researchers in the field, and to a lesser extent reading ethnographic studies about the field one wishes to study. The latter should perhaps be utilised as a supplementary stratagem rather than as the sole source of knowledge, as the more technical aspects of the environment may not be apparent. In particular, any acronyms, standardised forms or interview questions should be learnt, and this is best done either by asking or by observing in the particular setting where data will be collected.

The method I have used in this study is the second, building on prior knowledge, based on my experiences in healthcare- and advocacy-related work in the US and UK as well as as a competent participant in the geographic regions' cultures concerned. Primarily, my tasks have been to familiarise myself with spiritually oriented and humanistic psychotherapies. I have also had to familiarise myself with issues relating to recovery from addictions and neurological/mobility difficulties that become relevant in some of the psychotherapy sessions in my data set.

### *[7.2] Conveying Information*

Aside from the contextual (or sometimes ethnographic) knowledge often necessary to fully grasp the data of applied CA, is the data itself, by which I mean basic or pure<sup>11</sup> CA.

Conversation analysis, as mentioned previously, has a focus on the structural practices of primarily spoken interaction. It is sometimes called 'the study of talk-in-interaction' (Psathas 1994), although there has been considerable work on non-verbal and non-vocal

---

<sup>11</sup> Pure CA is a term that differentiates general conversation analytic research from applied conversation analytic research. In this instance, 'pure' is roughly equivalent to 'basic' in psychological research terms. It is possible that this difference in language is due to CA's roots in sociology, which is generally assumed to have an applied role in some capacity.

communication such as laughter (Holt, 2010), non-lexical utterances (Jefferson, 1978; Gardner, 1997), gaze (Goodwin, 1980), and gesture (Goodwin & Goodwin 1986) as well as interaction involving people with limited or no verbal capacity (e.g., Goodwin, 2003; Dickerson, Stribling & Rae, 2007).

Those who are unfamiliar with CA research are often baffled by detailed CA transcripts. To capture the details of interaction such as prosody, incomplete words, sound cut offs, and silences, conversation analysts use transcription conventions that are based on the innovations of Gail Jefferson (1983). Goodwin (1979) introduced specialised transcription for non-vocal behaviour, gaze in particular, and others have adapted this style to best represent their data and the features of interest. For example, Dickerson, Stribling & Rae (2007) study interactions between people diagnosed with aphasia or autism and their families, friends, carers, and teachers and often encounter complex overlap of participants' speech, body movements, and vocalisations and have developed transcription systems that can more easily portray this complexity.

The purpose of such detailed transcripts is to capture as much as possible of the interaction for two instances: working from a transcript as an aid to recorded data and to communicate as much as possible of the data in written publications. Features such as prosody and length of silences are continuously monitored in human interaction and distinguish meaning in nuanced ways. It helps the analyst to be able to see these differences on paper whilst working with the actual data (particularly when organising large numbers of similar instances into collections or during a data session), but it also makes possible understanding of the data in publications that would otherwise be based only on lexical, grammatical, and syntactical units. Transcription systems for non-vocal practices, especially gestures and gaze, are less well defined. Often authors will adapt a

system to best illustrate the phenomena under examination. Many of these systems are based loosely on that of Goodwin (1979), who first transcribed non-vocal practices in CA-informed research.

Today, many journals have online supplemental materials that can include audio or video files. However, access to these supplements may be restricted to subscribers and may not be accessible to users of older machines or business computers that restrict files types the user can access. In many articles, features such as timing of gaze or prosody are central topics (e.g., Goodwin 1979), and without transcription of these features, the author is unable to weave a convincing argument until a reader has the supplementary materials. This could undermine the quality of academic journals in that authors would not need to make their points in a succinct article but rather to rely on materials which are intended only to *supplement* that article. Additionally, many conversation analytic data have restrictions on their publication and authors must rely on transcription and possibly photo stills to convey the analysis. For example, a researcher may record physical examinations and not have permission from the participants to publish the visual component of the data despite having a finding that gaze is important in some aspect of those interactions.

In publications where readers are not assumed to be familiar with conversation analytic transcription, a glossary of transcription symbols is useful. Authors whose points do not require highly detailed transcription may opt to include minimal details and to state what those symbols mean. It is considered best practice to publish what is necessary rather than to overwhelm the reader with symbols that are not directly relevant to the analysis. What this means is entirely up to the author as well as dependent on the topic at hand. Because this is a doctoral thesis covering many features of interaction, it is necessary to use detailed transcription methods. I have prepared a glossary of transcription symbols

used in this thesis as well as examples of other transcription types which may be used in some of the references (Appendix A).

### [7.3] *Structures of Social Interaction*

Conversation analysis is not only rigorous but quite dynamic. It seeks to identify practices across an aggregate of data *and* to understand single cases based on those findings.

Similarly, it seeks to identify what practices do systematically within the organisations (or domains, as in Schegloff, 1996a) which will be discussed below *and* to explicate how those organisations come to bear on the particular moment in interaction. Schegloff (1996a, p. 437) describes these dialectics deceptively simply: ‘Conversation-analytic work plays back and forth between poles on several dimensions.’ The local implementation of non-case specific procedures and in turn the composite regularity of these procedures (thus making them non-case specific) is fundamental to the study of interaction through conversation analysis, both historically and methodologically.

A thorough and specialised vocabulary has developed around conversation analytic methodology. Description will always involve using terms participants would not themselves use. Although one approach to naturally occurring interactions might be to use colloquial terms or the terms that participants might use to describe phenomena, CA sacrifices participant orientation to the *language of the method* for participant orientation to the *structures present in their own interactions*. That is, CA tends to differentiate between phenomena that participants might call by one name (e.g., ‘laughter’; Potter & Hepburn, 2010), because participants actually treat the phenomena differently. CA is very much an observational method in that although CA takes a participant-focused approach, the focus is on how participants behave and demonstrate their understanding of, rather than how they make sense of, their own and others’ actions.

The most profound difference between CA and other observational methods is perhaps its treatment of the local and global. Whereas in other methods, ‘global’ may be society or the relationship and ‘local’ the behaviour in situ, CA treats the particular interaction, or interaction more generally, as global and the conversational structures in a *particular* interaction as local (Mick Finlay, personal communication, 2008). Although this description doesn’t constitute an official aspect of CA methodology, this treatment is evident in some of the language used by conversation analysts. For example, a locally initial person reference is one that is the first reference to a person in the immediately surrounding sequences rather than during the entire interaction or entire time the interlocutors have known each other.

There are multiple organisational structures present in conversation that may interact in a given piece of data. Schegloff (2007) has suggested six organisations: sequence, turn taking, word selection, trouble (or repair), action formation, and or overall structural organisation (or sequential organisation). All of these above organisations are highly interconnected and often in simultaneous operation. Although topic and preference can arguably be considered under sequence organisation, I suggest that due to their broader relevance these be considered separately and have done so in this chapter. However, we begin with sequential organisation, which is central to many conversation analytic observations and empirical studies.

### *[7.3.1] Sequential Organisation (Overall Structural Organisation)*

Ochs, Gonzales, and Jacoby (1996, p. 373) describe sequential organisation as a grammar for the production of talk-in-interaction. In the same volume, Schegloff, Ochs, and Thompson (1996, p. 32) discuss the co-organisation of sequential organisation. Actions are based on previous actions and create constraints for future actions (cf., Robinson, 2004, for



a treatment of the sequential context and functions of explicit apologies). In other words, talk is both context-shaped and context-renewing (Heritage 1984a, p. 242). Atkinson & Heritage (1984, p. 6) note:

no empirically occurring utterance occurs outside, or external to, some specific sequence. Whatever will be said will be said in some sequential context, and its illocutionary force will be determined by reference to what it accomplishes in relation to some sequentially prior utterance or set of utterances. As long as a state of talk prevails, there will be no escape or timeout from these considerations. And, insofar as unfolding sequences and their constituent turns are unavoidable analytic concerns for the interactants, they provide a powerful and readily accessible point of entry in the unavoidable contextedness of actual talk.

Related to local context is progressivity, the moving forward of a conversation and the completion of courses of action that gives way to new courses of action.

Openings/greetings give way to reasons of various sorts for the interaction, whether telling stories or asking for help, among many other possibilities and additional talk that makes up the bulk of the interaction. This body, to borrow a term from literature, gives way to closings/goodbyes – or in the context of a medical consultation, opening, presenting complaint, examination, diagnosis, treatment, and finally closing (Heritage & Maynard, 2006, p. 14). Within different stages of an interaction, different sequence types are most relevant. For an extreme example, one would not tend to find interlocutors saying goodbye before saying hello, unless perhaps jokingly. (For a more thorough treatment of openings, see Schegloff, 1968, 1979, 2004; and closings, Schegloff & Sacks, 1973.) Not only is overall progressivity an important feature of interaction, but so is more localised progressivity within the sequence. Sequences tend to be brought to completion with as

little extraneous talk as possible to impede its progression. Indeed, we find that insertions such as repair perform the task of ensuring intersubjectivity in order for the sequence to be brought to completion. For example, see Extract 1.

Extract 1 [TG 1:16–21 (Schegloff, 2007, p. 102)]

01 Bee: = [W h y ] whhat'sa mattuh with y-Yih sou[nd HA:PPY,] hh

02 Ava: [ Nothing.]

03 Ava: u- I sound ha:p[py?]

04 Bee: [Yee]uh.

05 (0.3)

06 Ava: No: ,

### [7.3.2] *Sequence Organisation*

Talk (and, as I will argue, certain non-vocal practices) is managed through sequences with beginnings and ends, and potentially pre-beginnings, post-ends, and middles, called pre-expansion, post-expansion, and insert expansion, respectively. Adjacency pairs, or base sequences (Schegloff, 2007), are these ‘beginnings’ and ‘ends’ upon which expansion can occur. Some examples of adjacency pairs that readily come to mind are question-answer (e.g., How was school today – It was okay) and greeting-greeting (e.g., Hello – Hi) sequences.

Sequences are generally discussed in terms of talk and turns (short for turns-at-talk). However, the courses of action embodied by sequences can be completed through non-vocal practices, including gesture but also the actual fulfilment of requests as shown by Lindströ (1997) and Schegloff (2007, p. 94). Berger and Rae (2010; under review/Chapter 8) show that communicative non-vocal practices, which have well known and precise meanings in a particular setting (such as nodding in response to a question in many English-speaking cultures), are treated as adequate responses under certain conditions,

particularly when there is the possibility of a problem being heard or when speech is impractical (e.g., during chewing or by someone who is otherwise unable to speak). Berger and Viney (2010; under review/Chapter 6) similarly show that nodding during television and film watching is treated as an adequate response.

### *[7.3.3] Action Formation*

To reiterate the practice-action relationship, practices are what our bodies (including our voices) do, and actions are what we do with those manifestations. Actions are what gets done and what brings about the relevance of other actions. Actions are brought into being by practices; they are the interactional meanings behind our behaviours. For example, saying 'hello' is a practice. The action embodied by that practice may be a greeting or a go ahead to speak (as with answering a telephone), among others. More often, practices are merely components of actions rather than the vehicle by which an entire action is carried out. For example, beginning a response to an enquiry with 'oh' or 'well' are regularly occurring practices, yet these words in themselves do not usually constitute the entire responsive action (Heritage, 1998; Schegloff, 2009). When they do, they are implicative of the kinds of verbal responses that characteristically follow them.

In large part, action formation is governed by sequence organisation and the courses of action that are embodied by sequences. Although sequences often embody particular actions such as invitations or requests, it is not always so clear from syntax and grammar alone. An invitation (the action) can be delivered as a question, as a statement, or even as a command (the form). The response tends to address the form before the action, so an invitation delivered as a question engenders not only the first pair part of an invitation-acceptance/refusal sequence but of a question-answer sequence as well. This brings us also to distinguish between actions and turns (cf. Levinson, 1983, p. 238). Whilst actions are

carried out in sequences of actions, turns are the spoken components of each action. Turns, or turns-at-talk, relate to the act of speaking and being heard, rather than to any specific actions. Because sequences of actions are often discussed in terms of turns rather than actions, it has become commonly accepted in CA that sequences consist of turns. Actions are commonly carried out through single turns, and turns commonly perform one action. However, this is not always (perhaps not even usually) the case, and having this in mind is very important to the study of non-vocal practices, which do not necessarily involve taking turns.

Speakers also use resources to express a particular stance, invoke a particular type action such as an invitation, and/or to convey other particular information at any given time. Such possible resources include all that can be done with the body, including speaking, and work together to convey a holistic meaning. For example, an offer that is delivered after a delay and with a sneer does something quite different from an offer that is delivered straight away and with a smile. So strong is our understanding that non-vocal practices can modify the meaning of our actions that the Rabbinic Sages said that one who gives charity with a harsh expression forfeits any merit that comes from giving the charity (*Shulchan Aruch, Yoreh De'ah, 249:3*). 'Harsh expression' in this instance refers specifically to facial gestures, although harshness in other modalities would be equally frowned upon.

Actions can be analytically specified regardless of the response they receive. Although misinterpretation or multiple interpretations are possible, this is what allows us to respond appropriately in situ as well as to do empirical work on interaction. Co-conversationalists continuously analyse current and preceeding actions and respond to them. They cannot see nor hear future actions to figure out the action of a preceeding turn.

Repair after next turn is also a demonstration that it is the speaker, not the recipient, who determines the intent of the turn and demonstrates intent in turn design and sequential placement. A speaker can of course decline to correct a recipient and thereby places alignment rather than intersubjectivity in the forefront. For example, in Extract 2, Al declines to correct Coral's understanding of his suggestion of specific melons as 'just different kinds of fruit'.

Extract 2 [CA 052109 Melon]

01 Al: Now we need some \*cantaloupe (1.1) some \*watermelon (1.5)  
 02 some \*honeyde:w ((\*gestures counting beginning with right  
 03 index finger, middle finger, and ring finger))  
 04 (3.4)  
 05 Co: Just different kinds of fruit  
 06 Al:>>((smiles, performs ring gesture/OK sign))  
 07 (0.8)  
 08 Al:>>°Yeah° ((turning away))

### [7.3.4] Turn Taking

Turn taking is the process by which we select who speaks next and allocate the length of turns-at-talk. Sacks et al. describe as a point of departure 14 characteristics of turn taking for conversation in their foundational 1974 work *A Simplest Systematics for the Organization of Turn-Taking for Conversation*. They then build a systematics that accounts for these 'grossly apparent facts' (Table 1), and alternate explanations are considered and refuted.

Sacks et al. provide a concise and broadly applicable account of turn-taking in conversation, from which the current section largely draws. Of prime importance is the turn constructional unit (ibid. TCU, number 13), because it is the smallest unit that

interactants have a right to produce before speaker change becomes relevant. Individual turns-at-talk are made of turn constructional units. They can be lexical, phrasal, or clausal, and conversation analysts are still working to further define other types of TCU. Table 2 shows the known types and examples of them.

Table 1. Fourteen grossly apparent facts (Sacks, et al., 1974, p. 700-701)
(1) Speaker-change recurs, or at least occurs
(2) Overwhelmingly, one party talks at a time <sup>12</sup>
(3) Occurrences of more than one speaker at a time are common, but brief
(4) Transitions (from one turn to a next) with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions
(5) Turn order is not fixed, but varies
(6) Turn size is not fixed, but varies
(7) Length of conversation is not specified in advance
(8) What parties say can vary
(9) Relative distribution of turns is not specified in advance
(10) Number of parties can vary
(11) Talk can be continuous or discontinuous
(12) Turn-allocation techniques are obviously used. A current speaker may select a next speaker (as when [s]he addresses a question to another party); or parties may self-select in starting to talk
(13) Various 'turn-constructual units' are employed; e.g., turns can be projectedly 'one word long', or they can be sentential in length
(14) Repair mechanisms exist for dealing with turn-taking errors and violations; e.g., if parties find themselves talking simultaneously, one will stop prematurely

<sup>12</sup> This observation is of particular relevance to silence. When no party speaks, both parties are not speaking. When one speaks, another is (usually) silent.

Table 2. Types of TCUs (Sacks et al., 1974)		
Type of TCU	Description	Examples
Lexical	one word or unit of utterance	Yes mmhmm
Phrasal	Not grammatically complete but can be understood as complete based on the grammar of the preceeding turn(s)	(Who?) Sarah's brother (Where?) On the roundabout
Clausal/Sentential	Grammatically complete on its own	Sarah's brother was driving on the roundabout

At the end of a TCU, a turn may be *possibly complete*. This means that the end of a TCU signals a point (the transition relevance place, or TRP) in which another speaker may begin speaking without being seen as interruptive. Interlocutors can go to great lengths to signal that a TCU or a turn is not complete through the use of perturbations, continuing prosody, repair (discussed later in this chapter), timing of silences to be at points when a TCU is not yet possibly complete, as well as gesture (e.g., holding a gesture across TCUs). Silences tend to occur within TCUs rather than following a possibly complete TCU, because within-TCU silences are simply pauses in which the speaker continues to have the right to speak (Sacks et al., 1974; cf. Hawkins, 1971)<sup>13</sup>. If the speaker wishes to continue speaking, they are likely to begin a new TCU immediately and then pause if necessary. Grammar is one of the strongest indicators of TCU completion, which makes projectability possible. Prosody also has strong effects on the projection of further talk vs. completion (Raymond, 2010).

<sup>13</sup> Although Hawkin's (1971) findings are consistent with Sacks et al. (1974) and follow a similar data collection method – recording spontaneous speech, Boomer's (1965) study with children giving elicited narratives finds that silences tend to occur between clauses rather than within them.



When a TCU is possibly complete, a next speaker may be selected either by the speaker of the current TCU, or a speaker may self-select (which can include the speaker of the current TCU producing another TCU) (Sacks et al., 1974). A next speaker may be selected by addressing the interlocutor directly (Sacks et al., 1974), directing one's gaze to another (Lerner, 2003), asking a question in which a particular interlocutor has epistemic authority (Lerner, 2003), or making a comment that shows one's own lack of knowledge in an area in which another has epistemic authority (Pomerantz, 1980). Sacks et al. provide a simple rule model for how next speakers are allocated (Table 3).

Table 3. Rules for selecting next speaker (Sacks et al., 1974, p. 704)
(1) For any turn, at the initial transition-relevance place of an initial turn constructional unit:
(a) If the turn-so-far is so constructed as to involve the use of a 'current speaker selects next' technique, then the party so selected has the right and is obligated to take next turn to speak; no others have such rights or obligations, and transfer occurs at that place.
(b) If the turn-so-far is so constructed as not to involve the use of a 'current speaker selects next' technique, then self-selection for next speakership may, but need not, be instituted; first starter acquires rights to a turn, and transfer occurs at that place.
(c) If the turn-so-far is so constructed as not to involve the use of a 'current speaker selects next' technique, then current speaker may, but need not continue, unless another self-selects.
(2) If, at the initial transition-relevance place of an initial turn-constructional unit, neither 1a nor 1b has operated, and, following the provision of 1c, current speaker has continued, then the rule-set a-c re-applies at the next transition-relevance place, and recursively at each next transition-relevance place, until transfer is effected.

Sometimes the selection of next speaker is by default the currently non-speaking party in a dyad. Although that person does not automatically speak, their non-response may be *relevantly absent*, particularly when an initiating action has high response relevance (such as a question). The current speaker can use *increments* (TCUs that grammatically fit with the prior turn but that are separated by silence or a strong prosodic end to the prior TCU) to continue the current turn or to give the appearance of continuing the turn in the event of a relevantly absent response or if a *dispreferred response* is forthcoming. A canonical example is shown in the extract below, in which Donny is calling Marcia to ask for a ride to work after having car trouble.

Extract 3 [MDE: Stalled (Schegloff, 1996b, p. 11)]

01 ((phone rings - once, possibly more than once))  
 02 Mar: Hello?  
 03 Don: 'lo Marcia,=  
 04 Mar: Yea[:h ]  
 05 Don: = [('t's) D]onny.  
 06 Mar: Hi Donny.  
 07 Don: Guess what. hh  
 08 Mar: What.  
 09 Don:>> .hh My ca:r is sta::lled.  
 10 >> (0.2)  
 11 >> 'n I'm up here in the Glen  
 12 Mar: Oh::.  
 13 (0.4)  
 14 Don: .hhh  
 15 Don: A:nd.hh  
 16 (0.2)  
 17 Don: I don' know if it's po:ssible, but {.hhh/(0.2)} see  
 18 I haveta open up the ba:nk.hh  
 19 >> (0.2)

20 Don:>> a:t uh: (.) in Brentwood? hh=  
 21 Mar: Yeah:- en I know you wan- (.) en I wou: (.) en I  
 22 would, but- except I've gotta leave in aybout five  
 23 min(h)utes. [(h)eh)  
 24 Don: [Okay then I gotta call somebody else. right away.

At line 9, Donny announces that his car is stalled, and at line 10, Marcia does not offer a response, for which the best case scenario would include an offer of assistance. At line 11, Donny produces an increment to line 9, and at line 12, Marcia produces 'Oh::.' which marks a change in knowledge. With no offer forthcoming, Donny makes moves toward making an explicit request at lines 17-18, to which Marcia again does not immediately respond (line 19). He produces another increment in the same fashion at line 20 and receives a refusal to his unspoken request. Such increments serve to give the appearance of a prior not having been complete and therefore that the silence is a pause within that turn rather than a gap in which the recipient should speak. They also allow the speaker to add information that may alter the recipient's response, such as proximity as in Extract 3.

### *[7.3.5] Preference*

Preference and its associated forms 'dispreferred' and 'preferred' are purely technical terms that do not necessarily reflect the intra-psychic preferences of participants. Extract 3 contains both an initiating action and a dispreferred responsive action in the request and refusal. Preferred actions are those that come readily and contribute to the progressivity of the course of action under way. With initiating actions, these are those that tend to come early in a conversation and are not mediated. In Extract 3, the purpose of the call was to get a ride to work and there was a sense of urgency. Although Donny got into the business of eliciting a ride to work quite early in the conversation, he did not immediately request it.

Instead he told of his troubles whilst progressing toward an overt request after Marcia does not respond with an offer. Indeed a request is never explicitly made. Requests are dispreferred initiating actions; offers are preferred to requests as shown in Extract 4.

Extract 4 [#17, ST (Schegloff, 1980, p. 112)]

01 R: 'Member the blouse you made a couple weeks ago?  
 02 L: Ya.  
 03 R: Well I want to wear it this weekend to Vegas but my mom's  
 04 buttonholer is broken.  
 05 L: Fred I told ya when I made the blouse I'd do the buttonholes.  
 06 R: Ya ((sniff)) but I hate ta impose.  
 07 L: No problem.  
 08 L: We can do them on Monday after work.

R has called L to request that she make buttonholes on a blouse. R does not ask directly but sets up the topic (lines 1-4). In line 5, L offers to make the buttonholes, and her offer is accepted (line 6). Extract 3 also contains a dispreferred responsive action, Marcia's refusal to pick Donny up, which in this instance is very intertwined with her withholding the offer to do so. This section of the extract is reproduced below for clarity as Extract 5.

Extract 5 [MDE: Stalled]

17 Don: I don' know if it's po:ssible, but {.hhh/(0.2)} see  
 18 I haveta open up the ba:nk.hh  
 19 >> (0.2)  
 20 Don:>>a:t uh: (.) in Brentwood? hh=  
 21 Mar: Yeah:- en I know you wan- (.) en I wou: (.) en I  
 22 would, but- except I've gotta leave in aybout five  
 23 min(h)utes. [(hheh)  
 24 Don: [Okay then I gotta call somebody else. right away.

At line 17, Donny begins to set the stage for an explicit request, giving Marcia the opportunity twice again to make an offer, but instead Marcia pre-emptively refuses. Marcia's refusal saves Donny from having to produce a dispreferred initiating action explicitly, but it is pre-emptive. However, Marcia's talk at lines 21-23 is delayed and only comes after Donny's latest increment at line 20. It is elaborate; it is not a simple 'No'. It is mediated and expressed as beyond her control; moreover, it accounts for not offering. Marcia says that she would [pick him up] except that she herself has to leave shortly. These features are also present in other refusals, such as Extract 6, in which A's elaborate response (lines 3-5) is delayed by perturbations and includes appreciation of the offer and a reason why she cannot accept.

Extract 6 [SBL T1/S1/C10, simplified]

01 B: And uh the- Uh if you'd care to come over and visit a  
 02 little while this morning, I'll give you a cup of coffee.  
 03 A:>> hehh! Well that's awfully sweet of you, I don't think I  
 04 can make it this morning, hh um I'm running an ad in the  
 05 paper and-and uh I have to stay near the phone.

Any time there is a silence following an initiating action, there is a greater possibility that the upcoming response will be dispreferred. Speakers of initiating actions may produce increments (Schegloff, 1996c) or rephrase the initiating actions in order to receive a preferred response, even if the new preferred response simply confirms the impossibility of the original initiating action. On the other hand, preferred responses usually come quickly and are much simpler. Compare these refusals to the following exchange in Extract 7 in which Nelson is calling Clara to ask her out for a drink.

Extract 7 [Schegloff 2007, JG 31, p. 30]

01 Nel: Y'wanna drink?  
 02 Cla:>> Yeah

Although most preferred responses do not get much elaboration, if any, they may lead to elaboration that progresses the course of action, such as making plans if an acceptance of an invitation. Dispreferred responses are those that are delayed or otherwise avoided, such as refusing an invitation. Dispreferred responses typically are accounted for or mitigated. For example, one might refuse an invitation by saying ‘Sorry, I can’t make it that day. I have to work.’, whereas one might accept an invitation by saying ‘That would be lovely!’. It would be exceedingly odd for a person to accept an invitation by saying ‘Yes. I have to work the day before, but I am off then. I’m terribly sorry I’m able to come.’ And quite rude to simply respond ‘No’.

Although types of actions are most commonly designated preferred or dispreferred, we speak of preference in terms of sequential organisation as well. For example, there is a strong preference for progressivity in talk (Stivers & Robinson, 2006). Such preferences have implications for other organisations. In particular, the preference for progressivity has implications for sequence and repair, with sequences unfolding to accomplish courses of action and repair having limitations on where it can be initiated without causing further trouble in the conversation. This will be discussed further in the next section.

#### [7.3.6] *Trouble/Repair*

Sometimes there is trouble in talk, and progressivity is compromised for the sake of intersubjectivity. This is called *repair*, and the target of the repair is called the *trouble source* or *repairable*. It is not necessary for the casual reader to grasp all of the technicalities of repair. It is useful, however, to know the following properties of repair (Schegloff et al., 1977):

1. Repair can occur at any point in the interaction.
2. Any utterance is a potential trouble source.

3. Repair can be initiated by any party. (Self or other initiation)
4. There is a preference for the speaker of the trouble source to provide the repair solution regardless of who initiates repair. (Self or other repair)

Self-initiated repair often occurs in the TCU in which the trouble source occurs or in the TRP following that TCU. These types of repairs are termed first and second position repairs, respectively. Self-initiated repair can also occur in subsequent turns, when a co-conversationalist has demonstrated a mishearing or misunderstanding, and is called third turn repair, in contrast with third position repair which is other-initiated. Other-initiated repair typically follows the turn that contains the trouble source and can range from an open class repair initiator such as ‘Huh’ or ‘What’ to seeking out specific details that would make the prior utterance understandable (e.g. What’s a USB cable?). It is rare for other-initiated (third position) repair to be solved by the initiating party and is in practice considered quite rude (e.g., ‘You mean article not manuscript.’).

Repair that is initiated after the next turn (i.e. third turn repair) is the last place where conversational structure provides for repair (Schegloff, 1992). Repair initiated much later in a conversation causes additional troubles in the talk and additional work is needed to pinpoint the trouble source. All such subsequent repair is termed fourth position repair and has not been studied extensively to date. A summary of where repair can occur (excluding fourth position) is provided in Table 4. I have used invented data here for illustrative consistency.

Table 4. Types of Repair			
	O- or S-Initiated	O- or S-Repair	Position
A: I went to the <b>ga-movies</b> .	Self	Self	First
A: I went to the <b>um</b> B: <u>The movies</u> ?	Self	Other	First
A: I went to the game <b>er movies</b> .	Self	Self	Second
A: I went to the game <b>I mean</b> B: <u>The movies</u> ?	Self	Other	Second
A: I went to the movies. B: <b>Where?</b> A: <u>To the movies</u> .	Other	Self	Third
A: I went to the game. B: <b>You mean movies</b> .	Other	Other	Third
A: I went to the movies. B: Who won? A: <b>I said movies</b> .	Self	Self	Third Turn
A: I went to the movies. B: Who won? A: <b>No</b> [ I said-] B: [Oh <u>movies</u> ] How was it?	Self	Other	Third Turn

### [7.3.7] Word Selection

The use of particular words instead of others can be significant – ‘Why that now?’

(Schegloff & Sacks, 1973) is perhaps the mantra of conversation analysts. Some work has been done regarding single words (e.g., ‘actually’, Clift, 2001). However, most word selection research has been in the area of person reference, whilst object reference has



often been assumed to follow similar patterns (Celia Kitzinger, personal communication 2010). In terms of person reference, the forms that are ‘usual’ in context are fairly well described. There is a preference to use the most familiar form known to the recipient and to use only one form and its associated pronouns. Locally initial references use the full form, and locally subsequent references use pronouns. Schegloff (1996a) refers to these uses as *reference simpliciter*, although some use the term *simply referring* (Celia Kitzinger, personal communication 2010). If a particular form is used in the place where it is expected, it is not doing anything special; it is reference simpliciter or simply referring. If ‘something is out of place’, we ask ‘Why that now?’ – What does that do at that moment in the interaction?

When a locally initial form is used in a locally subsequent position, it can do a number of things. For instance, it can distinguish the referent from others about whom interlocutors are talking, or it can topicalise the referent. Take for example the following exchange in Extract 8<sup>14</sup>, in which Sherrie’s use of ‘Alice’, in this case a locally initial reference form in a locally subsequent position.

Extract 8 [SN-4 Alice (Schegloff, 2007, p. 178)]

01 Mar: What about that girl 'e use tuh go with fer so long.

02 Kar: A:lice? I [don't-] they gave up.

03 Mar: [ (mm) ]

04 (0.4)

05 Mar: (°Oh?)

06 Kar: I dunno where she is but I-

07 (0.9)

08 Kar: ↑Talks about 'er evry so o:ften, but- I dunno where she is.

14 (0.5)

---

<sup>14</sup> I'd like to thank Celia Kitzinger for using this extract in her short course *Word Selection: Developing Skills in Conversation Analysis* and guiding our class in the analysis.

15 Mar:   hmh  
 16           (0.2)  
 17 She:>> Alice was stra::nge,  
 18           (0.3) ((rubbing sound))  
 19 Mar:   Very o:dd. She usetuh call herself a pro:stitute,='n I useteh-  
 20           (0.4) ask 'er if she wz gitting any more money than I:  
 21           was.(doing). (°An' she said-°) we'd compare notes yihknow.

Sherrie's use of 'Alice' here serves to launch a new topic, Alice's strangeness as well as to place Sherrie squarely in the conversation about her. Sometimes people do the opposite and use locally subsequent forms (predominately pronouns) in seemingly locally subsequent positions (which usually involve the full form) but in fact have not previously mentioned the people about whom they are talking. Sometimes they have mentioned a location or institution to which the referent belongs, but this is not always the case. Kitinger, Shaw, and Toerien (2012) recently looked at the issue of initial use of subsequent reference forms in English and found that this is a resource for establishing continuity of focus across sequences and conversations as well as for downplaying the relevance of a referent for the action being performed.

### *[7.3.8] Topic*

Students with whom I have worked have shown considerable intrigue when approaching topic organisation. Often topics have very permeable borders, with one point touching off another intermediate point, touching off another extended topic. In these cases, it is virtually impossible to pinpoint where one topic ends and another begins. This process may even occur during the space of a single turn-at-talk (Jefferson, 1984a). On occasion, topics are clear cut, with beginning markers such as 'Oh, that reminds me ...' or an abrupt topic shift following trouble. Topics can also be terminated through the use of idioms or summaries (Holt & Drew, 1995). Topics can additionally be organised around activities in

the environment or official processes. A medical interview's intake questions may be considered one topic, or more specifically a sequence of question-answer sequences. Viney and Berger (2010) have found that during television-watching, topics are very often arranged around what is on the screen or audible to the participants as in the following extract in which Kim and Barbara are watching the television show *The Apprentice*. Barbara is reminded by one of the contestants of an acquaintance, and this develops into talk about the acquaintance.

Extract 9 [Viney, KB 'Steve']

01           (1.6) ((Bar shakes head))  
 02 Bar:    >°mp I really don't like him°.<  
 03           (.)  
 04 Bar:    but he reminds me of (0.2) Steve. you know the one  
 05           who came to new ye:ars.  
 06           (1.6)  
 07 Kim:    no(h):, hh [heh heh]  
 08 Bar:                   [Lisa's ] fiancée.=  
 09 Kim:    =oh yeah:..  
 10           (0.4)  
 11 Bar:    (n- n- [e's real-)    ]  
 12 Kim:                   [why does he r]emind you of °Ste[:ve°.    ]  
 13           ((talk continues on Steve))

Topics may also be arranged around physical needs or disturbances, as in Extract 10, where Al's stray braid is topicalised in the form of joking.

Extract 10 [CA 052109]

01           ((Coral picks up glass, both look to television))  
 02 Al:       ((hair falls in eyes, shakes head  
 03 Cor:       ((gazes to Al))

04 Al: ((makes cutting motion with fingers, and growls))  
 05 Cor: hehhehheheheh that one (0.2) keeps fallin doesn't it darlin  
 06 hehehhehheh hh  
 07 (2.5) ((Al drinking))  
 08 Al: Sometimes I just want to cut it a:::::ll off ((making shaving  
 09 motion at head, continues until line 401))  
 10 (1.6)  
 11 Cor: Sure you do Al  
 12 (0.7)  
 13 Al: Be ba::ld ((moves hand around head))  
 14 Co: Sure you do (.) heh heh heh  
 15 ((talk continues on joking and Al's hair))

## **[8] Conversation Analysis with Silence and Non-vocal Practices**

Although accessible to interlocutors, CA research has (understandably) not looked specifically at non-vocal behaviours that do not have an overt communicative function – those performed peripheral to conversation (e.g., whittling). Rather, cooperative tasks such as setting up computers (Voilmy & Ting, 2007) and furniture construction (Rae and Guimarães, 2010) and the talk-based sequences used to accomplish them have been addressed in CA. The main focus of CA research on non-vocal behaviour, however, has been how people use non-vocal practices to augment (or sometimes replace) talk. Streeck (2009b), for example, describes in detail how hand gestures are used to enhance communication in German and English. Streeck maintains a cross-cultural focus on meaning and increasing intersubjectivity. Goodwin, on the other hand, describes resources that people with aphasia may use to communicate in the absence of competent speech (e.g., Goodwin, 2006). The non-vocal practices involved in Goodwin's work do not simply improve intersubjectivity but in many cases are the only ways in which intersubjectivity can be attained. Very few CA studies have addressed non-vocal practices that occur in sequential slots. Stivers (2008) demonstrates that nodding adds an element of affiliation to

recipient responses during storytelling. This practice, although it occurs responsively, does not occur in a position where a response is conditionally relevant. Chapters 7 and 8 of this thesis address the use of non-vocal practices in clearly demarcated sequential positions – first pair parts and second pair parts.

Silence has received more systematic attention in CA, with three types of silence named according to their position within or outside of units of talk – pauses within a TCU, gaps following a possibly complete TCU, and lapses following a complete sequence (Sacks, et al., 1974). Gaps are the most widely studied silences in CA, being a key sign of a forthcoming dispreferred response (Pomerantz, 1984a), and lapses are the least studied silences in CA due to the difficulty in obtaining recorded data that involve lapses of significant length. Most work involving silences has been oriented to establishing cross-cultural differences or similarities (Stivers et al. 2009; Stivers et al., 2010; Gardner et al., 2009; Mushin & Gardner, 2009), specifically with regard to tolerance for silence stemming from Jefferson's (1988) claim that approximately one second is a universal maximum silence in conversation. Chapter 3 adds to the debate about maximum silences and cross-cultural tolerance for silence. Another focus of silence-related CA research has been how pauses are used within turns to do specific work such as initiating self-repair (Schegloff et al., 1977) or setting up a topic as sensitive (Silverman and Peräkylä (1990; see also Silverman, 1996, pp. 63-88).

### **[9] Rationale for Using Conversation Analysis for this Project**

Silence and non-vocal practices (including gesture) are combined when non-vocal practices are performed without accompanying talk. Although non-vocal practices and silence are conceptually different, they can come together in such instances. Silence and non-vocal practices without accompanying talk are both treated as problematic in CA. I

take a methodologically based approach in this thesis to address this shared issue. It is my aim to explore how these interactional features are treated as unproblematic or are *used* by interlocutors to perform specific communicative functions that talk cannot. In this project, CA allows for identification over an aggregate and on a case-by-case basis of what functions practices perform. It allows for both in depth analysis of individual cases and analysis of a large collection of similar instances.

CA has been used throughout this thesis as the main method and theoretical approach. It is appropriate to approach methodological issues with the method under scrutiny as the primary method and to branch out as necessary to address particular concerns, for example that stem from aspects of the literature itself. This approach is most evident in Chapter 5, where I use quantitative content analysis and cluster analysis to explore how ‘incipient talk’ is used in the literature. A qualitative approach would not capture the links between different uses of this term, and a specifically conversation analytic approach would be clearly inappropriate for a textual analysis. Whilst a quantitative approach overall could provide a statistical account of particular words or coded types of questions across an aggregate, and seemingly giving this sort of dynamic, it takes out of context those particulars and imposes a theoretically based coding system on the data, unlike the inductive approach of CA. Some researchers have been using practices identified through CA as the basis for statistical analyses (e.g., Heritage, Robinson, Elliott, Beckett, & Wilkes, 2007; Stivers, 2007). This can aid in the dissemination of important findings to fields that more highly value quantitative methods, such as medicine.

CA does perhaps share some procedural aspects with grounded theory, in that both have stages of defining and narrowing scope based on the data. However, major differences between grounded theory and conversation analysis are the level of detail with

which single cases can be addressed and that the conversation analytic approach takes into account findings of previous conversation analytic research whilst allowing the data to inform the research. This is in contrast to both beginning with an intuitively or theoretically based coding system which is then revised based on the data or beginning with a clean slate each and every study. Phenomenology might be useful for a research question based on finding the essence of these interactions. However, it is not well suited for identifying conversational practices or resources that might not be consciously performed or remembered. Discourse analysis might be appropriate not only for a description of the interactions but what kinds of stances the participants take throughout an interaction. Again, it would not be well suited for the most central themes of this project. Conversation analysis has been selected due to its ability to work effectively over large data sets as well as with single cases and its suitability for the types of questions asked in this project. Whilst it is complex and specialised (and therefore requires extensive methodological training and continued practice), it is the most suitable for the task at hand. Prior to and during the course of this project, I have undertaken appropriate training both internally and externally and continue to do conversation analysis with a variety of clinical and non-clinical data in order to maintain and improve the quality of analyses. In the following chapter I will address in detail issues with classical treatments of silence in conversation, contextual issues in analysing silence, and methods for balancing accuracy of meaning and length of silences when discussing extracts.

### **Chapter 3**

## **Support and Evidence for Considering Local Contingencies in Studying and Transcribing Silence in Conversation<sup>15</sup>**

---

<sup>15</sup> A version of this chapter has been published as:  
Berger, I. (2011). Support and evidence for considering local contingencies in studying and transcribing  
silence in conversation. *Pragmatics*, 21(3), 291-306.



**Abstract**

Using a conversation analytic methodology, this chapter looks at conversations in English in which lengthy silences are regularly present. These silences are treated as unproblematic in this corpus. They apparently deviate from the proposals that gaps are minimized (Sacks et al., 1974) and that there is a standard maximum silence of one second (Jefferson, 1988). This is discussed in light of context and culture. Then the robustness of some features of the organisation of sequences (Schegloff, 2007) and turn taking (Sacks, et al., 1974) are considered. Finally, solutions are compared for rendering lengthy silences in such a way that their meaning is preserved in conversation analytic transcripts or other styles of transcription that include timed silences.

**[1] Introduction**

Sacks et al. (1974), in their foundational work on turn taking, show as one of fourteen 'grossly apparent facts' for which their model accounts, that silences and overlap are minimised in transition from one speaker to another, i.e. that no gap, no overlap between turns is a consequence of the systematics of turn taking. The time between turns is not presented as a steadfast rule of any specific time, but rather as a distribution with one beat of silence as the norm. This introduces the problematic of preference into turn taking organisation. Preference in conversation analytic literature refers to a pragmatically preferred action, rather than a psychologically preferred action. For example, the preferred response to an invitation is acceptance, not because of any desires on the part of the participants but as a consequence of the structure of conversation. Preferred responses tend to come without delay (silence, perturbations) and without much elaboration. On the other hand, dispreferred responses tend to be delayed and elaborate. See Schegloff (2007) for a detailed treatment of preference in sequence organisation.

After a speaker produces a first pair part, a delay in responding is often taken to indicate that a dispreferred second pair part might be forthcoming. When such a delay is silence, it can be termed a gap. (A response could also be delayed through the production of vocalisations such as uh, um, and so on.) Such gaps (or delays) may be an issue for the speaker of the first pair part. Since talk at this point is relevant and noticeably absent, speakers can do work, such as adding increments or rephrasing the first pair part, to elicit a preferred response (cf. Pomerantz 1984b). Gaps may be distinguished from pauses within a speaker's talk and from lapses, where talk (or vocal conduct such as laughter) does not occur and is not conditionally relevant. An example of the latter would be after a sequence has been completed (cf. Pomerantz 1984b).

Jefferson (1988), in an analysis presented as ‘preliminary notes’, suggests that speakers use a metric that results in speakers rarely allowing silences of more than approximately one second. She refers to this as a ‘standard maximum’ and shows here that longer silences tend to occur in the context of non-talk activities such as consulting a train timetable or writing directions. However, there have long been claims of cultural variation in silence in conversation (e.g., Goffman 1963: 103).

Stivers et al. (2009) tested the competing claims of universality and cultural variation by measuring inter-turn silence in 10 languages spoken in informal conversation. Although any similarities among 10 languages can hardly be called universal, possible differences can become known even with a relatively small data set. They found a difference of approximately 0.5 sec between the languages with the fastest and slowest mean responses (Japanese 7.29 ms and Danish 468.88 ms) and that the distribution patterns were similar across the 10 languages. Stivers et al. suggest that differences in response timing may be due to cultural differences in the *calibration* of what is considered delayed.

Whilst it is a worthy endeavour to test claims of variation or universality quantitatively, their study possesses potential problems with generalising beyond the specific languages and sequence types they examined. Their study was limited to polar questions, questions that are designed to receive ‘yes’ or ‘no’ responses. Because such questions have a relatively small pool of possible responses, this could result in more uniformity (and therefore lead to a false perception of universality). They base the generalisation from questions to all turn types across all languages on a corpus of Dutch data in which response times were not significantly different between questions and non-questions. It is worth noting that Dutch also did not have significant differences on response timing for confirmation/disconfirmation or gaze/no gaze, whereas most other

languages did. It is also unclear in what contexts the recordings were made. ‘Informal’ could mean anything from a couple talking over a meal to a lively gossip session among friends, and context could be a confounding variable making languages appear less similar than they may actually be.

Gardner, Fitzgerald & Mushin (2009) in studying interactions among Garrwa-speaking Indigenous Australian and English-speaking Anglo-Australian interactants, show that there was both a tendency toward lengthy silences (compared to canonical data) as well as an orientation to smooth speaker transitions as explicated in Sacks et al. (1974). Gardner et al. (2009) argue that tolerance for pauses and gaps is dependent on local interactional contingencies, namely a relaxed atmosphere and lack of orientation to ‘getting things done’. They maintain that whilst this attitude may be prevalent within a culture, a tendency toward long silences is not necessarily a property of the culture itself. This argument is supported by their analysis of Australian political debates in which there are quick responses to prior talk and indeed quite a bit of overlap. I will address these issues as they occur in non-Australian English data.

Sacks et al. used data that was primarily derived from audio recordings for both co-present and telephone conversations in California and the Northeastern United States in densely populated cities with white participants from a variety of socioeconomic backgrounds. The data presented here is from a variety of everyday and institutional interactions in large cities in England and the Mountain and Southeastern regions of the United States, with participants from white, Black<sup>16</sup>, and Jewish backgrounds. We do not propose that particular demographics necessarily account for differences, but rather that a diverse data set is necessary for understanding differences, or indeed similarities, among

---

<sup>16</sup> ‘Black’ is used here in recognition of participants’ ethnic identities rather than external categorisation denoted by ‘black’. It is used with the understanding that others may identify as ‘African American’ or use other terms. For further discussion, see McWhorter (2004), Gourley (1973), and *This Wicked Day* (2010).

interactions. Additionally, it should be noted that like the Gardner et al. corpora, the data presented in this chapter are video recordings of co-present interactions. Video recordings potentially allow analysts to attribute more accurate meaning to silences through knowledge of non-verbal behaviour, including gesture and engagement with the environment. These factors may affect whether a participant is taken to be not responding or in the process of responding, but there is no way to know whether they are occurring at all in audio-only data.

I shall provide additional evidence that tolerance for silence is shaped by local interactional contingencies, and indeed may shape or renew those contingencies. I shall show how interactants orient strongly to sequence organisation, speaker selection, and turn organisation despite this apparent deviation of long silences from canonical turn taking procedures. Finally, I will consider some of the issues in transcribing data that contains longer than canonical silences.

## **[2] Deviations and support**

### *[2.1] Deviation from canonical turn-taking features*

In the data we present, there are often long silences of varying lengths. Whilst this is not precluded by Sacks et al. (1974), they do treat the possibility of long silences as a tool for and feature of dispreferred responses and normative turn taking as no gap, no overlap. We find that this is not always the case for normative turn taking, and in fact the presence of silence within and between turns is often attributed to cultural differences in speaking and interaction (e.g., Tannen 1984).

The current research provides support for Gardner et al.'s argument 'against equating an instance of turn allocation that differs from that identified in the SSJ as a deviation tied

to a specific culture'. In other words, whilst there may be overall differences in speech rate or conversational rhythm across cultures and languages, this is not an absolute attribute across members of a culture or individual interactions within a culture or set of participants. I will demonstrate with a case of two different styles during the same interaction. The following two extracts are taken from the same psychotherapy session<sup>17</sup>, that is, the same therapist (Paula) and the same client (Leif) on the same day with the same speaker having epistemic authority and the floor, overall. Throughout this chapter, I will refer to extracts from Paula & Leif's sessions, primarily for uniformity for comparison between extracts but also because this subset of the data contains the lengthiest silences, and therefore the clearest examples, in the corpus.

These extracts are from video-recorded psychotherapy sessions of approximately one hour duration in a large but low population density city in the Mid-western United States. Sessions between Paula and Leif make up approximately six hours of data. They had been seeing each other for three years when data collection began, and the extracts presented are from their first recorded session. I loaned Paula a consumer-grade camcorder for the purpose of recording sessions, and both participants were able to direct when recordings were made and whether the recording of a session should be ended early or deleted. Paula is a humanistic psychotherapist whose training is in social work, and Leif is seeing her for issues related to both recovery from alcoholism and mobility problems. Both are middle-aged Americans, with Leif being somewhat older. They are from different parts of the Mid-western United States and have lived in their current area for several years. This data has been transcribed using Jeffersonian conventions, and silences were timed initially with

---

<sup>17</sup> There have been a number of articles on silence during psychotherapy sessions as part of a therapeutic approach (e.g., Prince 1997). Most of these have focused on the meaning of silences and are based on clinical experience rather than empirical studies. None have looked at silence in psychotherapy as a linguistic device. It does not appear to be the case that the silence discussed here is part of a therapeutic approach but rather is endogenous to the interaction itself. This is evinced by the presence of silence that can be attributed to either the therapist or the client regardless of prior speaker or who has rights or obligations to speak next. Additionally, there are instances of 'doing being silent' or 'doing not responding yet', which the majority of silences between this dyad are not.

a stop watch before being checked against the digital wave files extracted from the video.

Times for silences represent exact measurements rounded to the nearest 1/10 second. See

Section 3 for a discussion of silence transcription methods and relative vs. exact

measurement options.

In Extract 1, Leif is talking about himself and his day to day life, and it contains lengthy silences. He designs his turns so that silences are generally before a turn constructional unit (TCU) is possibly complete. Paula does not produce utterances to break these silences, treating them as pauses after which the TCU will be resumed. She does produce utterances after possibly complete TCUs and long silences separate their turns.

Extract 1 [PL 091609]

01 Lei: >Yeah it'd been kinda an interesting week=I haven't  
02 gotten this much accomplished< (1.6) as I'd like to  
03 this week  
04 (2.8)  
05 Lei: A:::n (1.1) I decided that maybe that's okay  
06 (1.6)  
07 Lei: I'm not gonna (0.7) beat myself up about it  
08 Lei: I mean nobody's (1.4) nobody ELSE'll beat me up  
09 about it so: (1.9) why should I be ↑first in li:ne?  
10 Pau: How long was it that you fired your boss?  
11 (1.0)  
12 Pau: °°(I mean)°°  
13 (1.7)  
14 Pau: How many months  
15 (2.3)  
16 Lei: Well January third so I guess that'd be:: uhh (1.6)  
17 you know (0.6) well goin on eight months  
18 (4.8)

19 Pau: 's still sinking in  
20 (0.6)

Compare this style to that of the next extract. In Extract 2, Leif is telling Paula how to solve a computer problem at the beginning of the session. His turns do not contain as many pauses, and silences are much shorter. There are even some instances of 'precision timing' in speaker transition.

Extract 2 [PL 091609]

01 Lei: So you bought it about a year ago  
02 Pau: Hmhm ((nodding))  
03 Lei: You probably have Leopard  
04 (0.9)  
05 Pau: ((nodding)) Y::es (0.7) I do  
06 Lei: =(It's an/Nice) operating system  
07 Lei: hh and and (0.6) Leopard (1.5) has the ability to  
08 recognise (1.2) a Windows file  
09 (0.8)  
10 Lei: When it sees one  
11 (0.7)  
12 Lei: It goes  
13 Pau: (okay)  
14 (0.7)  
15 Lei: A↑::hhh (1.6) that's one a THOse (.) files  
16 Lei: hh and so (0.8) a (0.1) a Windows (0.7) M S Word  
17 file (.) will open up (1.2) err-r show on the screen  
18 rather=  
19 Pau: =Mhm  
20 Lei: =(0.6) as (1.1) y'no a do::t do::c (1.7) file  
21 (0.5)



Overlap is virtually absent in the corpus. However, there does not seem to be an inverse relationship between lengths of silence and propensity for overlap. What little overlap we encounter tends to occur in contentious situations, such as in Extract 3.

Extract 3 [PL 101609]

01 Lei: No I don't miss [that  
02 Pau: [.hhuh  
03 Pau: You don't miss: (1.1) falling down becau::se of  
04 drinking  
05 Lei: W' I didn't usually do that  
06 Pau: Pardon  
07 Lei: I didn't usually fall down  
08 (0.8)  
09 Pau: Even with the PAWS  
10 (.)  
11 Pau: Even when um (0.8) your doctor was talking  
12 about the post withdrawa[l  
13 Lei: [That that really .hh ['s  
14 Pau: [That  
15 kind of was a time limi[ted thing  
16 Lei: [That was really ataxia  
17 (2.2)  
18 Lei: An so it was kind of a \*clugeon of\* of ((\*brings  
19 hands together interlocking fingers slightly  
20 during these words\*))  
21 Pau: ↓Yeah  
22 (1.2)

In this extract, Paula has enquired whether Leif misses drinking, which he has denied. She has challenged this with the argument that drinking is a lot like smoking (she is an ex-smoker), and that some times are perfect moments for a cigarette. She has gone on to list a

few things that she proposes he doesn't miss. This extract immediately follows that list.

Paula now (at line 3) asserts that Leif doesn't miss falling down *because* of drinking. This is something that Leif has apparently not experienced, and he is quick to correct the apparent assumption. Paula tries to show her knowledge of previous events by introducing withdrawal as a time when Leif fell down (i.e. not because he was drunk, as previously implied). In lines 12-15, there is significant overlap as they compete to establish the correct interpretation of Leif's situation. In lines 11-12, Paula has rephrased her question from line 9. This ends with terminal overlap from Leif's turn at line 13. Leif's turn is then in terminal overlap with Paula's next turn at line 14, which suggests that his falling down was time limited. This turn is then recognitionally overlapped in the middle of 'limited' with Leif's next turn at line 16, 'That was really ataxia'. There is silence for 2.2 seconds, and Leif produces a further explanation of his prior situation during the time to which Paula has referred, that it was a 'clugeon of ((gesture))'<sup>18</sup>.

A gloss of this turn is that the situation was a bundle of different problems put together. At this point, Paula concedes to Leif's epistemic authority about his own medical problem. Although it may appear that Paula is attempting to write Leif's history for him, she has gotten into a tight interactional place with her earlier assertion (as well as the assumption that Leif must miss drinking sometimes) and must choose between explicitly acknowledging that she was wrong or attempting to find some basis for her assertion after the fact. Nevertheless, the participants must establish what will be accepted as fact in the local interactional environment (regardless of what they may actually believe).

---

<sup>18</sup> 'Clugeon' is a regional (American Mid-west and Southeast) word that is not generally written and is used to mean a somewhat chaotic mixture or a tight jumble. Leif's gesture is that of mashing, which reflects the definition. I have been unable to locate a standard spelling and have transcribed it here to resemble 'bludgeon', which has a similar pronunciation.

In the previous extracts, we did not encounter any clearly dispreferred responses. In Extract 4, Paula enquires as to how Leif is doing (line 2). For 3.5 seconds he maintains his body and head position that he had prior to Paula's question and then looks up and says 'I'm okay' and smiles (line 4). This initial response is delayed and is analysable as disingenuous (thus being dispreferred in structure albeit with a word choice that would otherwise be a preferred response). Note that the delay is not necessarily in timing alone. Leif holds his position during the gap at line 3. He is *doing* not responding. Paula treats 'I'm okay' as disingenuous and challenges Leif's assertion by asking him to account for his reasoning, which after two attempts results in Leif confirming that he is indeed not okay. He has significantly delayed his dispreferred response in the form of beginning with linguistic elements of a preferred response but with the prosody expected in a dispreferred response.

Extract 4 [PL 091609]

01 Lei: ((holding hands together and gazing down))  
 02 Pau: So how are ya  
 03 >> (3.5) ((Leif holding position))  
 04 Lei:>> ((looks up)) °I'm okay° ((smiles))  
 05 (1.1)  
 06 Pau: How do you know  
 07 (2.4)  
 08 Lei: I said so:  
 09 (2.3)  
 10 Pau: On your own authority you are okay  
 11 Lei: No  
 12 (1.9)  
 13 Lei: ((sniffs))

In this extract, there are some very long gaps. However, there are equally long gaps in other extracts in which there are no signs of dispreference, for example Extracts 5 and 6.

Extract 5 [PL 091609]

01 Pau: >What did you do<

02 (1.4)

03 Pau: <I know it was magic

04 >> (3.9)

05 Lei: Uh::m (2.8) I (1.1) saved (0.6) th' files (3.1) and

06 renamed them

07 >> (2.8)

08 Lei: For instance (0.6) uh (2.1) uh- (4.2) th- (0.3)

09 >first one was called uhh< (0.8) oh gee I drew a

10 blank uhh reservation er something like that

Extract 6 [PL 091609]

01 Lei: Well January third so I guess that'd be:: uhh (1.6)

02 you know (0.6) well goin on eight months

03 >> (4.8)

04 Pau: 's still sinking in

05 (0.6)

06 Pau: There's no one to account to

07 (0.6)

08 Pau: <↑But your wife

09 (1.5)

10 Pau: An the dawg

11 >> (2.2)

12 Lei: ↓Yeah

Communication (formerly Speech) Accommodation Theory (Street & Giles 1982;  
Giles & Wiemann 1987) as well as Bateson's schismogenesis (1935) state that people

change their communication styles to suit the context, and specifically to be in line with their interlocutors' styles. In interactions where hostility and power are exhibited (e.g., cross-examinations, Eades, 2007), there may be very little accommodation. Differences can be treated as defiance on the part of the witness. Yet in political debates (another potentially volatile interaction but where participants have more equal status), parties adjust their volume and speech rate and have considerably more overlap (including interruptions) in the service of having their views heard instead of the opponents'. Whilst it is beyond the scope of this chapter to consider these theories in detail or to conclude why or at what point interlocutors adjust their styles, this field of linguistic knowledge provides further evidence for the treatment of silences as contextual rather than directly cultural. Cross-cultural studies on silence in similar interactional situations could further clarify the roles that cultural and situational contexts play in the presence, use, and treatment of silence in human interaction.

*[2.2] Orientation to the larger organisations of turns and sequences*

Much of the silence in this data is unmarked and unproblematic. By this, I mean that participants do not treat the silence as signalling a forthcoming dispreferred response, as non-uptake, as a possible hearing problem, or as a possible understanding problem. In the case of the former two, the first speaker may produce increments or rephrase the first pair part in the service of receiving a preferred response. In the latter two, the first speaker may initiate self-repair. In all of these instances, the silence is accountable, whether the second speaker's silence is due to actively doing being silent or to the first speaker's difficulty communicating. Extracts 7 and 8 show examples of marked or problematic silences.

Extract 7 [PL 091609]

01 Pau: .hhhooo

02 Lei: =Should I smile a lot

03 >> (0.5)

04 Pau: Pardon

05 >> (0.5)

06 Pau: Sure

Extract 8 [PL 091609]

01 Pau: You're still the dinner (0.6) cooker aren't you

02 >> (1.1)

03 Pau: (I mean) at your house?

04 >> (2.1)

05 Pau:>> Maybe not on (0.5) on uh (0.2) >Thursday nights

06 >> (0.9)

07 Lei: On Monday and Thursday nights there seems to be a

08 reproachment

With unmarked and unproblematic silence, on the other hand, participants show no orientation to such issues but rather continue the expected actions. The differences in what is treated as unmarked and unproblematic silence between the canonical data and that which I (as well as Gardner et al.) present could lead one to the conclusion that the interlocutors are not orienting to turn-taking or sequence organisation. However, the presence of silence and the resumption of talk are strongly organised around both turn taking and sequence organisation:

A. Silences before the possible completion of a TCU (pauses) tend to result in the same speaker continuing or restarting the TCU that was in progress before the pause.

Extract 9 [PL 091609]

01 Lei: S<sub>1</sub>o:: instead of- (0.5) calling it by its full name  
02 I just called it recrui:t (0.6) do:t (1.0) do::c

Extract 10 [PL 091609]

01 Lei: [Well either either or because it was coming (1.8)  
02 that way it probably (0.7) it probably didn't (1.9)

Extract 11 [PL 091609]

01 Pau: hh now remember that (0.4) ih- I don't have a TV  
02 thing tuh (0.4) show us through it (0.6) >to look at

In all three of these extracts, silences occur at points in which the TCU has not been brought to possible completion, and the current speaker continues following the silence. In Extracts 10 (line 2) and 11 (line 2), we also have silence before first position self-initiated repair, 'it probably (0.7) it probably didn't' and 'show us through it (0.6) >to look at'. This pattern shows orientation to the points at which turns come to possible completion as primarily based on grammar and syntax and that speakers have a right to bring turns to possible completion before another speaker takes the floor.

B. Silences after the possible completion of a TCU to which a responsive action is relevant but before a sequence has been closed (gaps) tend to result in the prior speaker producing an increment or the recipient producing a sequentially relevant response.

Extract 12 [PL 091609]

01 Pau: Otherwise I have to get a little (1.1) TV °n I  
02 don't really have one°  
03 (0.6)  
04 Lei: Television  
05 >> (0.5)

06 Pau: Yeah  
 07 >> (1.3)  
 08 Lei: £That's okay£

In Extract 12, there are two gaps at points in which there is a relevant response. Line 5, following Leif's other-initiated repair, shows a gap of 0.5 second that is ended by Paula responding affirmatively (line 6). Following Paula's confirmation, there is a 1.3 second gap that is ended by Leif's 'that's okay'. This thus closes the sequence that was begun by Paula regarding the availability of a television for viewing the tapes that they are recording.

Extract 13 [PL 091609]

01 Pau: You're still the dinner (0.6) cooker aren't you  
 02 (1.1)  
 03 Pau:>> (I mean) at your house?  
 04 (2.1)  
 05 Pau:>> Maybe not on (0.5) on uh (0.2) >Thursday nights  
 06 (0.9)  
 07 Lei: On Monday and Thursday nights there seems to be a  
 08 reproachment

Extract 13 is the same as Extract 8 that we have previously seen. In this extract, Paula is enquiring about Leif's cooking responsibilities to his family. Line 1, the original enquiry, does not receive a response from Leif before Paula speaks again (line 3), adding detail to line 1. Then after 2.1 seconds, she produces another turn, giving an alternative at line 5. This finally receives a relevant response from Leif at line 7. This pattern, when compared to treatment of pauses, shows orientation to the possible completion of a TCU in which the initiating action is done. It also demonstrates orientation to the sequential relevance of initiating and responsive actions, including repair insert sequences.



C. Silences after the possible completion of a TCU and after a sequence has been closed  
(lapses) may result in either speaker beginning a new sequence.

Extract 14 [PL 101609]

01 Pau:    Zoning out in front of the TV  
02 Lei:    Habitual sex uh (.) I don't know  
03        >> (0.9)  
04 Lei:>> hh but [uh  
05 Pau:                [((clears throat)) scuse me  
06        >> (3.0)  
07 Lei:>> I I I've been thinking um hmmneh a lot about .hh  
08        that last month or so I've also been having a lot  
09        of sleep in my eyes the last couple three weeks

In Extract 14, the participants have been talking about alternatives to drinking in a humorous tone. Lines 1 and 2 are the end of this topic, and line 3 is a 0.9 second lapse. Leif ends this lapse apparently to begin a new topic that is cut off by Paula having to clear her throat (lines 4-5). After Paula has cleared her throat, there is a 3 second lapse (line 6), and Leif moves to begin a new topic at line 7.

Extract 15 [PL 101609]

01 Pau:    [What about it was so seductive  
02        (0.9)  
03 Lei:    .hhhh Wl I've been thinking about that,  
04        (0.8)  
05 Lei:    †I really have  
06        (1.6)  
07 Lei:    Um  
08        (1.8)  
09 Lei:    MOSTly what I didn- (3.3) mostly what I don't miss

10           (2.1) is disappointing myself  
 11           >> (0.7)  
 12 Pau:>> Hm

Extract 15 follows from Extract 3. In the current extract, Paula has asked (line 1) what was so seductive about drinking. In lines 2-10, Leif formulates and states his answer (and is seen to be doing thinking during the silences). After Leif's answer, there is a 0.7 second lapse (line 11), which Paula ends with 'hm' at line 12. The topic is then, several seconds and another 'hm' later, picked up again by Leif (data not shown).

This pattern shows orientation to sequence closure as well as the points at which a new sequence may begin. Whether there is in fact a lapse or a new sequence, rather than expansion of the sequence, is beyond the scope of this chapter and is indeterminable for the participants at the point in which there is silence. Also, the presence of lapses is certainly more common in co-present data in which the main activity is attention to a stimulus or physical action rather than talking. Note that in the psychotherapy data in this corpus, along with a tendency for long gaps and pauses, lapses tend to be of similar length or even shorter than gaps and pauses.

Rather than not orienting to turn-taking and sequence organisation or having an entirely different systematic, they are orienting so strongly that intra-turn silences (A) are treated as pauses rather than abandoned TCUs; inter-turn silences where a relevant response should be produced (B) end in pursuit, second pair part, or repair; and inter-turn silences where a sequence has closed (C) are treated as appropriate places for either party to begin a new sequence (or engage in post-expansion). This is strong evidence in support of the robustness of the organisation of turn allocation and turn design (Sacks et al., 1974) as well as the organisation of sequences (Schegloff 2007).

A detailed analysis of the meaning of silences is beyond the scope of this chapter (but see Chapter 4 on one use in psychotherapy). Certainly there are cases in which silences are in the service of ‘doing not responding’ or ‘doing thinking’ (as in Extract 15). I have chosen to focus here instead on the unproblematic nature of silences that would in many data sets indicate problems or markedness in their very length. I now move into a discussion of the difficulties and possible solutions for transcribing such interactions without giving the appearance of interactional troubles and without sacrificing precision or accuracy.

### **[3] Considerations for Transcription**

Video data gives analysts information about what is actually happening during silences of any length as well as during speech or other sounds. The non-vocal actions present in audio data may be overlooked (as in the case of gestures before speaking, which could be wrongly interpreted as a delayed response) or inferred (as in the case of giving an item or opening a window). In either of these cases, researchers working with only audio data may not have a complete picture of the interaction. In working with video data, different styles of transcription and level of detail may be appropriate for non-vocal activities depending on the focus of the research. There are currently many variations on how gesture and non-vocal behaviour are transcribed, most of which are variations on Goodwin’s (1979) basic approach (e.g., Streeck 2002; Stribling, Rae, & Dickerson 2007) and may be supplemented with pictorial representations or photographs (e.g., Goodwin 2002). The wide variety of research that involves non-vocal behaviour necessitates transcription styles suited particularly for the research at hand. For this reason, I have chosen to focus on the transcription of silence itself, which can be done in combination with an appropriate transcription style for the non-vocal behaviour that is relevant to the analysis.

One difficulty in transcribing data with slower than canonically described speech or longer than canonical silences is that what is in effect a barely noticeable silence in context can be quite long when compared with canonical data. Analysts who are familiar with canonical data may interpret such transcribed silences as doing things they are not. For instance, quite a short silence (a micropause, canonically less than 0.1 second) in some of the data presented is roughly 0.4 second and a typical pause or gap lies in the range of 0.7 to 2.0 seconds. Whilst the participants do not treat these as long silences, analysts who are accustomed to working with more canonical data may, particularly when working with only transcripts, as one often does when reading a journal article or book chapter. Psathas and Anderson (1990) state that the relative length of a given silence compared to the syllable length and silence between words (components of speech rate) is more important than the actual length of silence. There are a number of options that a transcriber can use to convey the data accurately in time and meaning, although a balance must be struck according to the focus of the research. I will now outline some of the options available and discuss the benefits and drawbacks of each.

*[3.1] Giving approximate values to silences by using a fixed rating system*

This approach can be subdivided into three methods.

A. Defined levels

e.g., less than 0.5 sec, 0.5-1.5 sec, and above 1.5 sec represented as (-), (- -), and (- - -)

This method allows for data with fairly discrete differences but with unequal intervals to be characterised as short, medium, and long silences (or as many levels as appropriate).

B. Defined levels with specification

e.g., less than 0.5 sec, 0.5-1.5 sec, and specifying lengths over 1.5 sec represented as (-), (- -), and (4.7)

This method allows for data with fairly discrete differences but with potentially very long silences to be represented compactly.

C. Increments of 0.5 sec

e.g., less than 0.5 sec, 0.5-1.0 sec, 1.0-1.5 sec, etc. represented as (-), (- -), (- - -), etc.

This method is similar to the 0.1 sec based system that is widely used in conversation analytic transcripts when symbols are used instead of numeric values to represent lengths of silences. However, it allows for a less precise interval. For very long silences, a + can be used at 5 seconds, much the way it is used at 1 second in the traditional 0.1 sec based method.

These methods attempt to capture the approximate length of a silence relative to itself and not the surrounding talk. They simplify the usual CA transcription of silences by making the interval of interest larger and therefore less precise. However, care must be taken to develop intervals that are meaningful in light of the data. There may be quite a difference in how a silence at the beginning of an interval and a silence at the end of an interval are treated by the participants. These methods can also be more difficult to understand, when the transcriptionist uses symbols that are widely known to mean other measurements. I have used dashes here as a matter of convenience and aesthetic. As long as the transcription conventions are made clear in a given manuscript, one simply needs to consider the adjustment that readers must make to different meanings or new symbols.

*[3.2] Using a conversation-based metric for timing silences*

This method was pioneered by Wilson & Wilson (2005) who, drawing on Wilson & Zimmerman (1986)'s experiments on the regular rate that varies between conversations with which turn-taking options are recycled in the SSJ model, argue that neuro-oscillations in each of the participants that are kept in synchrony via the rate of syllable production are

responsible for such regularity and indeed is necessary for coherent conversation. They argue that when participants speak slowly, i.e. the rate of syllable production is low, both silences and sound production are stretched and that there should not be long silences in moderate or quick speech. Wilson & Wilson (2005) go on to suggest that silence should be transcribed such that each beat of silence is relative to the surrounding talk. For example, (0.7) would represent 7 relative beats of silence and a micropause, as the participants treat it but regardless of length, as (.), with of course a note about the alteration to the transcription convention. At present there is no precise metric by which to make these judgements, but it is in development.

This chapter does not argue for the use of a speech rate metric for timing silences at present. Although this may be appropriate for some conversations and participants, there are challenges that this corpus presents to such an approach:

- A. Participants' speech production rate varies across the data within a conversation and even within turns.
- B. Lengthy silences are present regardless of the rate of syllable production and are not treated as problematic. Although they may be lessened in some contexts, these shorter or less frequent silences are still not in line with canonical examples<sup>19</sup>.
- C. This theory does not account for coherent conversations between parties whose syllable production rates differ (although note that people tend to adjust their speaking style to more match their interlocutors, as mentioned in Section 2.1).

---

<sup>19</sup> This pattern is also present in Finnish speech, with syllable rate being 305/min compared to 291/min in English, yet silences in Finnish accounted for a similar amount of time as other European languages (Lehtonen & Sajavaar, 1985, p. 197).

*[3.3] Taking the mean silence between words during smoothly flowing talk and subtracting this from timed silences.*

This entails timing the silences as one would normally when doing conversation analytic transcription as well as the silences between utterances during smooth spates of talk. The mean of the silences during smooth spates of talk is then subtracted from every recorded silence to give what is expected to be the real ‘silence’ as recognised by the participants. This can be a very tedious process and is not necessarily an accurate depiction despite being precise.

*[3.4] Timing silences as they are and conveying in a description of the material that longer than canonical silences are typical of the interaction.*

This method retains the standard format of CA transcripts and is therefore easy for an analyst to read without keeping in mind new symbols or new uses of existing symbols. It does not convey in the transcript the nature of the silences in relation to other talk or their typicality, but instead asks the reader to bear in mind that lengthy silences are nothing out of the ordinary and do not necessarily mark a forthcoming dispreferred response.

#### **[4] Conclusion**

I have shown that lengthy silences can be present in conversation as an unproblematic and regularly occurring feature. I have also shown that despite its regularity, it is subject to local interactional contingencies and may be present in the same environment as precision speaker transition and even overlap. I have focused on a part of the corpus that has particularly long silences in order to compare situational and structural differences between the same participants. Rather than generalising the results of this study to all American English or all therapeutic situations, the objective of this chapter has been to demonstrate that although early conversation analytic research as well as Jefferson’s (1988) proposed

metric was focused on American English, there are contexts in which the standard maximum silence of one second does not hold. This is in line with findings in Australian English and Garrwa. Moreover, even though lengthy silences are present, there is a strong orientation to turn allocation and sequence organisation.

Taking into account that some comparatively long silences may not mean the same as silences of similar length in many other English transcripts, I have addressed the problem of transcribing such silences both accurately and meaningfully and offered some solutions that readers may adapt to their own data. I have argued for or against some approaches more strongly than others. However, by no means should these solutions be taken as better than any other, or better than each other, overall. Each has its strengths and weaknesses and is more appropriate for some data than others. Readers may even find that different styles prove better for showing different phenomena. In this thesis, I take approach 3.4, reporting silences accurately in terms of length and conveying that extended length does not automatically mean a problem with the interaction or turn. Although this chapter dealt with the possibility of contextual differences in the interpretation of silence, it addressed how the length of silence before a dispreferred response might differ across contexts. However, it did not address contexts in which silence might be an indicator of a forthcoming or in-progress preferred response. The next chapter explores this issue in the context of psychotherapy and shows that silence is one way in which clients can perform sincerity in responding to the therapist's talk.



## **Chapter 4**

### **How Silence Contributes to the Performance of Sincerity in Psychotherapy<sup>20</sup>**

---

<sup>20</sup> The version presented in this chapter has received minimal comments from my supervisor. A version of this chapter is currently being revised for publication and will include John Rae as a co-author.

### **Abstract**

There have been a number of articles on silence as part of psychotherapeutic approaches (e.g., Prince, 1997). Most of these have focused on the meaning of silences, particularly from psychoanalytic perspectives, and are based on clinical experience rather than empirical studies. These authors tend to interpret silence as pathology or resistance in the client and/or advise clinicians to be especially tolerant of silence and to avoid filling silences. Cook (1964), on the other hand, found that sessions with 4-20% silence (counting only silences over 5 seconds) were rated as most successful, indicating that silence is not necessarily a form a resistance, as many believe. Hill, Thompson, and Ladany (2003) point out that nonetheless the correlation of silence with outcome ignores the placement of silence and the psychological needs of the client. No research has looked at silence in psychotherapy as a linguistic device to manage the local sequential environment, however. This chapter will use conversation analysis and draw on findings from everyday and clinical interactions to explore this issue.

Silence following an initiating action is characteristic of dispreferred responses such as refusals and negatively valenced answers in general conversation (Sacks, Schegloff, and Jefferson, 1974; Schegloff, 2007). However, this is not the only instance in which silence following an initiating action occurs. Silverman (1996, p. 66; see also Silverman & Peräkylä, 1990) found that silence was one way in which people would set up topics as sensitive in HIV/AIDS counselling sessions. The data from which this analysis is drawn is 9 spiritually oriented humanistic psychotherapy sessions. This chapter considers how delayed responding is one way in which sincerity is conveyed by clients.

**[1] Introduction**

In the previous chapter, I examined contextual differences in tolerance for silence and how silence does not always indicate that a dispreferred response is forthcoming. There may, however, also be instances in which silence not only is unproblematic but is treated as necessary for a preferred response to be treated as such. I examine this possibility in the context of psychotherapy, where there have been a number of published articles on silence as part of a therapeutic approach. Most of these have focused on the meaning of silences, particularly from psychoanalytic perspectives, and are based on clinical experience rather than empirical studies. Some explicitly advise clinicians to be especially tolerant of silence and to avoid filling silences. Most publications dealing with silence in psychotherapy are written from prescriptive, theoretical positions by practitioners who are writing from their own interpretations of their own practice.

Many writings on silence in psychotherapy are written from psychoanalytic perspectives in which the therapist is taught to allow silence to occur regardless of possible relevance of speaking (e.g., Prince, 1997). In psychoanalysis, silence is taken to be meaningful and attributable to the client with the exception of when the psychoanalyst decides the client is ready for an interpretation (Haim, 1990). Blos (1972) claims that attributions of the meaning of silences can be made based on the psychoanalyst's feelings about the silence. Fliess (1949) even argues that silence is a form of sphincter closure to cause retention (of words, substituting for excretory products). Such retention is supposedly intended to punish the psychoanalyst and is an act of transference of feelings about a dominant and cold parent to the psychoanalyst. The idea that silence can be equated to client aggression has been a particularly salient feature of psychoanalysis. The *Journal of the American Psychoanalytic Association* published a special issue in 1961 on silence as a form of acting out (cited in Ephratt, 2008). Gale and Sanchez (2005) challenge

these traditional views of silence, operating from within a psychoanalytic framework and arguing that taking silence out of the realm of language leads to mistakes in interpretation and denying clients ‘quiet areas’ within the therapeutic community (cf. Davies, 2007, who reviews in detail the ways in which silence has been characterised in psychoanalytic theory).

Surprisingly, no research has looked in depth at silence in psychotherapy as a linguistic device involved in managing the local sequential environment – beyond the simple presence of ‘interactional silence’ (e.g., Frankel, Levitt, Murray, Greenberg, & Angus, 2006). Most empirical research in this area has instead measured the talk time of each party or examined proposed intra-psychic meanings of silence through direct interpretation or interviews (e.g., Levitt, 1998, 2002). However, talk time and silence are not direct opposites, as talk can occur in overlap and each party is usually silent during another’s talk (Jefferson, 1984b, 1986; Sacks et al., 1974). Mahl (1956) uses the formula ‘Patient-Silence Quotient = N Seconds of Silence/N Seconds Available to Patient to Talk’ yet characterises such silences as dysfluent. Others measure the presence of silence or patterns of silence but with little reference to interactional function (e.g., Hill, 1978; Tomicic & Martínez Guzmán, 2011). Importantly though, silence is not monolithic in its placement nor meanings (Sacks et al., 1974; Levitt, 1998; Hill, Thompson, & Ladany, 2003). This chapter will consider how silence, particularly in the form of delayed responses to the therapist’s talk, is one way in which sincerity is conveyed by clients.

Although speaker transition regularly occurs with relatively little silence in conversation (Sacks et al., 1974), longer silences following an initiating action are characteristic of dispreferred responses such as refusals and negatively valenced answers in general (Schegloff, 2007). However, this is not the only instance in which silence

following an initiating action occurs. Silverman (1996, p. 66; see also Silverman & Peräkylä, 1990) found that silence was one way in which people would set up topics as sensitive in HIV/AIDS counselling sessions. Coupland and Coupland (1997) likewise observed silence as a feature of death-implicative talk (e.g., discussion of organ failure) in geriatric medical consultations. In all three cases, silence is used to show sensitivity to social expectations or norms, including that of demonstrating sensitivity to the emotional state of recipients.

Cook (1964) presents one of the few empirical studies on silence in a strictly psychotherapeutic context. He used five ‘successful’ and five ‘unsuccessful’ (as rated by the therapist; these ratings were not used in the analysis) cases from a previous study (Tomlinson & Hart, 1962) that had tested a scale unrelated to silence – the Process Scale (Walker, Rablin, & Rogers, 1960). From each session, Cook extracted nine two minute segments which were then rated for successfulness using the Process Scale. A ‘client’s silence ratio’ was calculated for each of the segments. Only silences of five seconds or more and that were terminated by the client were counted toward the total proportion of silence (the ‘client’s silence ratio’). As with most other authors that have explored silence in psychotherapy, Cook makes the assumptions that silences are attributable to the client and that the ‘owner’ of the silence will be the one to break it. A more robust approach would have been to look at silences that occurred after the therapist was the last speaker. Nevertheless, Cook found that segments with the highest Process Scale ratings (as rated by two judges) had 4-20% silence. Interestingly, segments that had more than 20% silence were rated as more successful than segments with less than 4% silence. The current study explores one reason why this might be the case, namely whether quick responses are treated as less sincere or as otherwise problematic by the therapist than delayed responses.

The data from which this analysis is consists of 9 hours of spiritually oriented humanistic psychotherapy between a middle-aged woman, Paula, whose psychotherapeutic training is in social work, and two separate clients. It does not appear to be the case that the silence discussed here is part of a therapeutic approach but rather is endogenous to the *interaction*. This is evinced by the presence of silence that can be attributed to either the therapist or the client regardless of prior speaker or who has rights or obligations to speak next. Additionally, there are instances of ‘doing being silent’ or ‘doing not responding yet’, which the majority of silences between these dyads are not. The silences that I examine in this chapter are those that occur in sequential positions in which the silence is attributable to the client, who is tasked with providing a therapeutically relevant response to the therapist’s prior action.

Silence in the psychotherapy context occurs between two or more people. Silence in this context inherently comes about as a result of nobody speaking and therefore everybody engaging in silence – what McDermott and Tylbor (1983) term *collusion*. Despite this, silence in interaction is often attributable to a particular party or parties, making a verbal response relevantly absent. However, silence does not always simply mean absence. Although a response might be absent or not *yet* produced, silence creates anticipation in parties that were not implicated in responding. This anticipation can lead to additional, intervening talk or more attentive postures. The ‘silent’ party (i.e. the implicated party) might also engage in an assortment of postures and facial expressions to demonstrate that a response is forthcoming, often called ‘doing thinking’. Even when ‘nothing’ is happening, the implicated party can be *doing* nothing. Silence is therefore *performative* and *communicative* in sometimes subtle ways. In co-present interaction, silence is inevitably accompanied by non-vocal behaviour, even if that behaviour is to remain still. By adopting a posture, one communicates ‘doing nothing’, ‘thinking’, etc.

through one's silence. These behaviours provide the context for others to understand the attributable silence – whether the initiating action has been heard and understood and whether a response is forthcoming.

Although as observers, we are unable to definitively say whether a particular silence is due to intra-psychic factors, there is at least an interactional claim that such silences are due to various intra-psychic difficulties that may be exploited in interaction. Chafe (1980) discusses several factors that influence the presence of silence in conversation. Some of these include pragmatic boundaries and transitions, whilst at other times silence indicates a difficulty in transition, demonstration of a change of state (cf. 'oh' as a change of state marker, Heritage, 1984b), or searching for a word (cf. Goodwin, 1983).

One issue with studying silence in psychotherapeutic contexts is consideration of the roles that the therapist(s) and client(s) take. Baker (1955) found that when psychiatrists took a passive role, clients were more engaged and had a vocabulary expansion rate similar to ordinary conversation. However, when psychiatrists adopted a more active role, clients were put into a defensive role and had less variety in their speech. This is a similar phenomenon to 'machine gun questions' reported by Tannen (1985) in which the recipient may feel obliged to respond quickly, with precisely what was asked, and in the minimum language required to do so. Goodman (Esterly, 1979, cited in Tannen, 1985) describes the effect that a fast-talking style such as this can have on those unaccustomed to it – upset, dissatisfied, incompetent, though they may not understand why – and postulates sources of the fast-talking style – anxiety, domination, boredom, the need to express freshly stimulated thoughts. He describes this style as crowding, and it is obvious why psychotherapists might not want to engage in this style and might see clients who engage in it as problematic. Although the adoption of a fast-talking style is not necessarily due to

intra-psychic problems, incongruence between institutional and personal communication styles can cause severe misunderstandings.

In a different situation that often involves considerable threat to face and potentially dire consequences – interrogation of witnesses in courtrooms – silence is often advised in order to communicate thought-through and consistent responses, with some manuals advising no less than a 5 second gap following the question (Summit, 1978). However, opposing lawyers frequently interpret silence as lying or resistance (Walker, 1985). Cross-culturally, interpretation of silence can have detrimental effects. Eades (2000, 2007; cf. Mushin & Gardner, 2009) found that Indigenous Australians whose backgrounds include regularly leaving long silences before verbally answering questions were frequently misinterpreted in court as uncooperative. This is likely compounded by ethnic stereotypes but is interactionally based in the silence. The lawyers in this study did not recognise that many Indigenous Australians' answers to questions actually include considerable silence, i.e. as preparation and not simply gap. Gardner, Fitzgerald, and Mushin (2009) propose that this situation engenders a 'clash not of fundamental cultural turn-taking practices, but of tolerance for silence in a particular institutional setting.' Berger's (2011/Chapter 3) work on co-present US and UK data supports Gardner's idea that a lack of orientation to 'getting things done' contributes to silence in conversation and that rather than having culturally bound turn-taking practices, silence can vary situationally within a culture.

Using the same psychotherapeutic data from which Berger (2011/Chapter 3) draws, I will explore how the conversational style and therapeutic agenda sometimes clash in this way such that conversationally appropriate (lack of) silence is treated by the therapist as resistance or insincerity, whereas conversely, the practice of leaving gaps contributes to the performance of sincerity. In order to examine the nature of silence in psychotherapeutic



contexts, two studies were conducted using quantitative and qualitative methods. Study 1 examined the data for the presence of gaps following four types of initiating action commonly found in the data and conversation analytically examined cases in which clients produce responses immediately following the therapist's therapeutically relevant turn. Study 2 used CA to examine in detail two minutes of a therapy session in which silence (and the lack thereof) was key to the breakdown and realignment of the interaction.

## **[2] The Data**

The data from which the current analysis is drawn involves Paula, a spiritually oriented humanistic psychotherapist, and two separate clients, Sally and Leif. All participants are middle-aged Americans of European descent from the western United States. Paula is a woman who is trained as a clinical social worker, and Sally is a woman who was assigned male at birth and has been seeing Paula for assistance with her gender affirmation for approximately two years. Her sessions comprise approximately three hours of the corpus. Leif is a man who has been seeing Paula for guidance in his recovery from alcoholism and has mobility problems that cause him considerable anxiety. He has been seeing Paula for approximately three years. His sessions comprise approximately 6 hours of the corpus.

## **[3] Immediacy is Dispreferred: Accounting for Immediate Responses to Therapeutic Questions**

One approach to identifying preferred responses within a sequence type is to examine the frequency with which different practices occur in responsive positions. Gap is one of the best indicators that a dispreferred response is forthcoming, so ordinarily one would expect gaps following initiating actions to be uncommon. This was, however, not the case for overtly therapeutic questions. Out of 130 therapeutic questions, 120 (92.3%) were followed by gaps. Ten instances (7.7%) across the nine sessions involved therapeutic questions that

were responded to immediately. These will be examined here in light of the local sequential environment. Immediately, Extract 1 can be identified as delayed responses despite an immediate utterance.

Extract 1 [PL 100109]

01 Pau: And how do you know it's depressh: depressed

02 Lei:>> I don:-

03 Pau: W'I [I mean

04 Lei:>> [I

05 Pau: Instead of pleasant

06 (1.6)

07 Pau: What's the quality of it

08 (.)

09 Lei:>> .I do(h)n't feel good about it

10 (2.4)

09 Pau: What are you thinking

In Extract 1, Paula's self-repair intervenes during Leif's production of an answer. In order to maintain the floor, Paula is engaging in rather quick uptake of turns. Leif finally produces his full response in line 8 following twice beginning what could be the same answer ('I don't', 'I', and finally 'I don't feel good about it'). Following Leif's immediate response (line 9) to Paula's final question (line 7, 'what's the quality of it'), Paula indicates that his answer is inadequately specific by asking for more information (line 9). In Extract 1, the delays were caused by Paula's self-repair, rather than Leif doing sincerity or thoughtfulness, and Paula treated Leif's otherwise immediate response as inadequate. This leaves nine instances of immediate responses to therapeutic questions for which accounts can aid in understanding how participants orient to the interactional requirements of the therapeutic agenda.

*[3.1] Structurally Provided-for Responses*

The initiating action in Extract 2 is actually the first pair part of an insert sequence pursuing a response following a long gap of 10.4 seconds (line 12). Paula has given an interpretation (lines 1-11) to which Leif has not yet responded. He responds immediately (line 14) to Paula's pursuit in line 13. In this instance, Leif has already exhibited delay, and further delay might demonstrate a lack of engagement, rather than thoughtfulness. Recall that Cook (1964) found that there was an optimal proportion of silence in his client-centred psychotherapy data.

Extract 2 [PL 010810]

01 Pau: You're the same person three weeks ago  
 02 (1.2)  
 03 Pau: And yet (2.4) your ability to (0.1) see the situation  
 04 differently (1.8) each time you encounter it (2.8) is  
 05 greater than it used to be  
 06 (4.4)  
 07 Lei: ((cough))  
 08 Pau: [It's like you're not (.) thinking habitually  
 09 (1.2)  
 10 Lei: ((cough))  
 11 Pau: Your approach to things is not so: (2.8) stereotypic °anymore°  
 12 (10.4)  
 13 Pau:>>How do you like that  
 14 Lei:>>Well I like it fine  
 15 (4.3)  
 16 Lei: ((cough)) Of course naturally (2.1) I've tried to  
 17 figure out why

Although a response is not conditionally relevant following a statement, Leif makes moves to engage by placing coughs between Paula's turns and (not shown) using body

movements typical of taking the floor. Following Paula's interpretation, he abandons these practices and demonstrates that he is thinking. Paula's question builds on implicit acceptance of her interpretation in that 'how do you like that' presumes that Leif agrees that her interpretation is accurate. Her question also makes a response conditionally relevant and directs Leif in forming a response by providing a question-answer framework that may be easier for him to articulate due to its specificity. It also focuses him on the therapeutic agenda by directing him in the kind of information she is seeking.

Also following an interpretation, Paula asks Leif a question about one aspect of the problem that is directly related to her interpretation (Extract 3). He has not spoken during the interpretation, and Paula's turn ends with a yes/no tag question, to which Leif responds affirmatively and immediately. Stivers & Rossano (2010b) found that tag questions are one way to mobilise responses when a response is not conditionally relevant from a sequence organisational perspective.

Extract 3 [PL 100109]

01 Pau: hh so what is it (.) and if it passes (0.4) is there a  
02 way for you to relate to it (1.6) differently than you  
03 >> do now when it comes because there's kind of an it  
04 >> quality to it isn't there  
05 Lei:>> Yeah  
06 Pau: You know you're not saying to me (1.6) I didn't get my  
07 bike out and I meant to all weekend and it was a little  
08 windy so I didn't do it and then I got depressed

Extracts 1-3 involve interpretations and confirmation seeking in the form of questions. In ordinary conversation, confirmation seeking prefers yes responses, which are routinely presented without delay. However, without elaboration on the part of the client, the

therapist treats these seemingly preferred responses as inadequate. In Extract 1, Paula explicitly asks for elaboration, whilst in Extract 2, she remains silent for 4.3 seconds before Leif self-selects to elaborate. In Extract 3, Paula elaborates on her own prior turn, accounting for the interpretation by citing what Leif has (not) said, thereby treating Leif's turn 'yeah' as a dispreferred response to her interpretation.

### *[3.2] Immediate Responses to 'Obvious' Questions*

Again, Paula presents an interpretation (Extract 4). This time, she presents Sally with an interpretation using rising intonation (line 4), marking it as a question to which a response is conditionally relevant. Sally confirms this interpretation (line 5) with an oh-prefaced response (Heritage, 1998), marking it as an obvious answer for which the question should not have been asked. Agreement is the preferred response in this position, both due to Paula's turn design and the nature of the sequence. In this case, Sally treats Paula's prior action as inappropriate and displays this through her oh-prefaced response. Following this sequence, a 1.4 second lapse develops before Sally makes an 'um' vocalisation and pauses for 2.3 seconds before changing the topic to a piece of internet art that someone e-mailed her.

Extract 4 [PS 122309]

01 Sal: And we (2.3) had uh oh shaggy bells instead of silver silver  
 02 bells heh  
 03 (2.4)  
 04 Pau:>> So there is play once in a great while  
 05 Sal:>> Oh yeah  
 06 (1.4)  
 07 Sal: Um (2.3) somebody e-mailed me an (1.2) engineering drawing on  
 08 how to (0.9) uh () how to erect a Christmas tree

In Extract 5, Paula produces a preliminary to a therapeutic question (lines 1-4), in which she rephrases her query-in-progress as a reflexive wondering. Leif does not respond to this, and following a 4.4 second gap (line 5), she accounts for her earlier turn with 'I mean I don't know that just occurred', followed by an overtly therapeutic question 'so do you have any theory on why you don't get out on the bike if it's a little too hot'. Leif responds immediately to this with 'yes' in line 9. Following a 1.9 second gap (line 10), he reiterates and expands on his 'yes' response and produces a piece of paper (lines 11-12). This paper contains the weekly news letter from Alcoholics Anonymous and is a regular feature of their sessions, to which Paula orients by laughing in line 13. In line 14, Leif indicates that Paula has forgotten something from their last session, 'we were just talking about that the other day weren't we' with emphasis on 'weren't we', thereby marking her question as having an obvious answer (line 14). This ultimately results in Paula having to check her notes in line 17. Kitzinger (2011) has elsewhere documented interactional trouble when interactants pretend to remember and are later caught not remembering. In this extract, Paula is caught not remembering, rather than pretending to remember. There is, in any case, trouble in the interaction because of it. Paula asked a question that made visible her not remembering, which is then treated by Leif as having an obvious answer and receives a dispreferred response.

Extract 5 [PL 012210]

- 01 Pau: Have you ever wondered if it just occurred to me (0.2)  
02 I wonder how often our excuses (0.8) reflect our fears  
03 you know we make the excuse before we're even conscious  
04 of what's the (.) fear that going and doing this  
05 (4.4)  
06 Pau: I mean I don't know that just occurred  
07 Pau:>> So do you have any theory on why (1.0) you don't get out on  
08 the bike if it's a little too hot

09 Lei:>> Ye:::s.  
10 (1.9)  
11 Lei: Yes I think it has a lot to do with it I uh as a matter  
12 of fact ((takes out paper))  
13 Pau: heheheh  
14 Lei:>> We were just talking about that the other day weren't we::  
15 Pau: °Yea:h°  
16 Lei: I think we were  
17 Pau:>> I think you were what were we talking about ((checks notes))

Sometimes participants treat questions as having obvious answers that are not evident specifically from previous turns or previous interactions, but rather from broader contexts. Extract 6 involves Sally describing a neurosurgical procedure which she does not want to have. She is talking about being awake during brain surgery as being one of the top reasons she doesn't want to do it (lines 1-4). Paula then asks 'and that really unnerves you' (line 5), which Sally treats as obvious by using an immediate, oh-prefaced response (Heritage, 1998) in line 6.

Extract 6 [PS 110409]

01 Sal: But you know they have to (1.8) stimulate things in  
02 (1.6) you know what happened (1.1) do you get back  
03 memories from the past er (0.8) do you smell  
04 something er (1.6) °does it affect motor function°  
05 Pau:>> And that really unnerves you  
06 Sal:>> °Oh yeah°  
07 Pau: Do you suppose that's something that you and I can  
08 prepare you for

*[3.3] Disaffiliative Immediate Responses*

Disaffiliation is one possible feature of dispreferred responses to initiating actions.

Disaffiliation is when a party displays a stance that opposes that of another party in the interaction (cf. Stivers, 2008). This may be done explicitly or implicitly, as in Extract 7. In this extract, Sally has been discussing how her neurologist has been pressuring her to enter a clinical trial for deep brain stimulation, a neurosurgical treatment for Parkinson's Disease that involves implantation of a pacemaker into the brain and that, according to Sally, only temporarily 'resets the clock' on the disease and involves significant risks that she is not willing to take.

Extract 7 [PS 110409]

01 Pau:>> Have you ever talked to someone who's had it

02 Sal:>> Yes

03 Pau: Okay more than a few people

04 (2.7)

05 Sal: UM (3.5) I guess that depends on how close you're (6.0) h(h)how

06 how one to one the talk has to be

07 (1.2)

08 Sal: Y'know I've as far as presentations in front of groups=

09 Pau: Mm

10 Sal: =um (2.3) I've seen y'know probly a dozen or so

Paula asks Sally whether she has talked to someone who has had the procedure (line 1), to which Sally responds immediately 'yes' (line 2). Although Paula's turn in line 1 is designed for a yes response, which Sally provides, Paula treats 'yes' as inadequate and requests more information (line 3). Following this request, Sally produces longer turns with large amounts of silences, both gaps and pauses, and accounts for her 'yes' response by expanding on her sources of information. Paula now treats this response (lines 5-10) as adequate by using the minimal response token 'mm' (see Gardner, 1997, 2002).



Extract 8 involves Paula asking Leif a follow-up question to how Leif feels about making excuses. Leif has said that he wonders what his grandmother would say if he was making excuses around her, and Paula and Leif have just completed a repair sequence related to the unexpectedness of Leif's comment (data not shown). This extract is on the same general topic as Extract 5 and follows it. Paula's question (lines 1-2) 'What would your granma if she was here right now what would she be saying' is a therapeutic question designed to have Leif articulate his thoughts on excuses. Leif responds immediately and ironically with 'she probably wouldn't say anything she's probably whack me up the side of the head with a frying pan and tell me to get out and start pickin potatoes or something' (lines 5-6). Paula does not, however, treat Leif's turn as a joke. She instead responds with an interpretation 'so her response would be don't over don't over analyse it'.

Extract 8 [PL 100109]

01 Pau:>> [What would your granma (0.7) if she was here right now what  
02 would she be saying  
03 Lei:>> Sh[e probably w]ouldn' say anything she's probably=  
04 Pau: [( )]  
05 Lei:>> =whack me up the side of the head with a frying pan and tell  
06 >>me to- (1.2) get out and start pickin potatoes or something  
07 (0.7)  
08 Lei: Pulli°n' potatoes°  
09 (0.9)  
10 Pau: S:o (0.3) her (0.3) response (1.2) would be (1.1) don't over  
11 don't overanalyse it

Soon after, in Extract 9, Paula attempts a therapeutic intervention to help Leif be less judgmental to himself (lines 1-6), in which she invokes the image of his grandmother hitting him with a frying pan. Leif responds immediately in line 7 with a laxly pronounced affirmative/negative token 'nyeah', which can be produced to do sensitive interactional

work (Jefferson, 1978). In this instance, Leif is agreeing with the intervention but disagreeing with the image through which his self-judgment has been conveyed. Because Paula did not treat his previous turn as a joke, there is a chance that she has understood the image as a realistic one. As we see in his next turn at lines 8 and 10, Leif takes back his previous image of his grandmother (Extract 8, lines 5-6). As Leif is beginning this turn, Paula begins simultaneously 'It's just you know', orienting to the negative element of 'nyeah' and displaying accountability for the interpretation. Paula drops out, however, and does not articulate a full account. Instead, she produces acknowledgement tokens 'mkay' (line 11) following possible completion and 'mmhm' (line 12) following a second possible completion of Leif's turn. Ultimately, Paula implicitly accounts for the intervention in lines 15-17 by responding that although Leif's grandmother might not be violent, he might be (figuratively) violent to himself.

Extract 9 [PL 100109]

01 Pau: >We enjOY IT but hhh so I guess what I'm e-askin is (0.3) could  
 02 >>you take the grandmother:: (0.2) image (1.2) and instead of it  
 03 >>being (1.2) someone whacking you with a ruler or a frying pan  
 04 >> and saying >eh get back to work< (0.8) judging you (1.3) what  
 05 >> if she's an aspect of yourself that knows: (1.1) this too shall  
 06 >> pa:ss  
 07 Lei:>> Nyeah  
 08 Lei:>> [Actually y']know she she was she really wasn't the type=  
 09 Pau: [It's just you know]  
 10 Lei:>> =that'd take a frying pan or[ a] roller to somebody  
 11 Pau: [°°mkay°°]  
 12 Pau: Mmhm  
 13 Lei: But uh  
 14 (1.4)  
 15 Pau: But you might be

16 (1.3)

17 Pau: To yourself

The final case (Extract 10) requires a lengthier discussion and benefits from greater background information. Therefore, rather than providing this final extract here, I will now present an analysis of the entire first two minutes of the session from which this extract has been taken. These two minutes include an interaction that has gone awry and the eventual realignment of the parties' talk.

#### **[4] Silence is Preferred: An In-Depth Analysis**

I have taken a segment of the beginning of a psychotherapy session between Leif and Paula. This particular segment has been selected due to inclusion of all of the key points in the analysis: a characteristic delayed 'sincere' or 'thoughtful' response, a quick response that minimises, and a quick response that challenges. I will present an analysis of this segment with a focus on the use of silence or lack thereof by Leif, the client, and how the associated utterances are treated by Paula, the therapist.

##### *[4.1] Characteristic Delayed 'Sincere' or 'Thoughtful' Responses*

I would like to begin with an example that exemplifies a 'sincere' or 'thoughtful' response to the therapist's talk. Extract 10 involves Paula and Leif discussing Leif's former drinking.

Extract 10 [PL 101609]

1 Pau: [What about it was so seductive

2 (0.9)

3 Lei: .hhhh Wl I've been thinking about that,

4 (0.8)

5 Lei: †I really have

6 (1.6)

7    Lei:        Um  
 8                (1.8)  
 9    Lei:        MOSTly what I didn- (3.3) mostly what I don't miss  
 10                (2.1) is disappointing myself  
 11                (0.7)  
 12   Pau:        Hm  
 13                (2.0)  
 14   Pau:        Hm  
 15                (1.0)  
 16   Lei:        In other words (0.5) spending ti::me (6.6)=  
 17   Pau:        hmhehhu[h ((cough))  
 18   Lei:        =[basically getting wa:sted (2.3) when I probably  
 19                could have been doing s(h)omething else  
 20                (.)  
 21   Pau:        °Yeah°

In line 1, Paula produces a request for information 'what about it was so seductive'. Leif immediately displays thoughtfulness by remaining silent briefly (line 2) and reporting that he has been thinking about Paula's very question on his own (lines 3 and 5). Line 3 is followed by another brief silence, and line 5 is followed by a lengthier one of 1.6 seconds (line 6), 'um' (line 7), and another lengthy silence of 1.8 seconds (line 8). Leif begins to formulate his actual response in line 9 'mostly what I didn-', followed by a 3.3 second pause and continuing with a replacement repair 'mostly what I don't miss', a 2.1 second pause, and 'is disappointing myself'. Leif's silences before he begins his actual response demonstrate consideration of his response and are broken up to emphasise that a response is in the works and thus that Paula should not pursue a response. Leif is then silent for another approximately 4 seconds (lines 11-15), during which Paula displays affiliation with his response though two 'hm's (lines 12 and 14). Leif has not yet provided an in depth answer, although he has provided what appears to be a sincere answer. Paula's refraining

from talk at this point shows recognition that there is more to come or that more is expected, whilst her 'hm's show that Leif is on the right track in terms of what kind of answer is required. In line 16, Leif continues by expanding his prior turn with 'in other words (0.5) spending time' followed by a 6.6 second pause. Paula coughs (line 17), and Leif continues his turn (line 18) with 'basically getting wasted'. This phrase ends with continuing intonation and has been transcribed as continuous with the following talk despite the 2.3 second silence before 'when I probably could have been doing something else' (lines 18-19). Paula's 'yeah' (line 21) explicitly affiliates with Leif's response and serves as an acknowledgement token.

Extract 10 involves an apparently sincere response to a therapeutic question by the therapist in which the client uses lengthy silences. Although relatively long silences are common with these participants, this extract contains very long silences up to 6.6 seconds. In the subsequent sections, I will show how the presence of lengthy silences of this sort are not mere happenstance but are one way in which clients can demonstrate sincerity in the psychotherapy session. I will contrast the smooth flow and silences of Extract 10 with that of the immediately preceding discussion to show how responses with lack of adequate gap following interventions are treated as incorrect or insincere responses by the therapist and how quick responses can be used by the client to challenge or minimise the therapist's preceding action.

#### *[4.1] Uses of Quick Responses*

##### *[4.1.1] Quick Responses that Minimise*

Extract 11 is the beginning of the exchange that includes Extract 10. It involves an instance of a quick response that minimises the therapist's prior talk. Immediately prior to Extract 11, Paula has finished setting up the camera in order to record their session for the day, and

the pair have exchanged a brief, unrelated sequence. Paula's topic initiator 'do you miss it' (line 1) does not have a clear referent. Leif's 'what' (line 3) constitutes an other-initiated repair on 'it', and Paula's 'drinkin' (line 5) is the repair solution. Lines 2, 4, and 6 contain silences; however the typical micropause or inter-turn silence for this dyad is approximately 0.4 seconds. These silences do not on their own constitute 'silence' for the participants, as they do not ordinarily orient to silences of this length as problematic (See Chapter 3). Thus Leif's 'no' (line 7) comes in quite quickly. He follows this up at line 9 with 'not really'. Paula does not attempt a turn during this time nor for another 0.7 seconds (line 10). Thus she has not treated Leif's 'no' as an adequate response.

Extract 11a [PL 101609]

1	Pau:	hh do you miss it
2		(0.6)
3	Lei:	What
4		(0.5)
5	Pau:	Drinkin
6		(0.5)
7	Lei:	No
8		(0.4)
9	Lei:	Not really
10		(0.7)

When Paula does speak again (line 11), she challenges Leif's assertion that he does not miss drinking. Leif does not respond during the 1.1 seconds (line 12) of silence following Paula's 'really' (line 11). Paula produces an account for her disbelief (lines 13-14) comparing drinking to smoking (she is an ex-smoker). After a 0.8 second gap (line 15), she expands this turn with an emotive account of how one can miss smoking (lines 16-17), followed by a 0.9 second gap. She further continues, bringing her talk back to Leif 'you don't miss (1.9) you don't miss drinking' (line 19) and 'you don't miss time went on' (line

20). This receives no response from Leif over a 3.8 second gap (line 21). Paula then produces a direct suggestion of what Leif doesn't miss about drinking 'you don't miss Patty [Leif's wife] getting mad at you'. After a brief gap, Leif agrees (line 24), and they both laugh (lines 24-25).

Extract 11b [PL 101609]

11 Pau: Really?  
 12 (1.1)  
 13 Pau: Cause I thought for most people it's sorta like  
 14 (0.9) s:moking (.) you miss tha:t  
 15 (0.8)  
 16 Pau: There are moments when you just (0.8) boy it'd be  
 17 a perfect moment for a cigarette you know  
 18 (0.9)  
 19 Pau: You don't miss: (1.9) °you don't miss drinking°  
 20 Pau: You don't miss time went on  
 21 (3.8)  
 22 Pau: You don't miss Patty getting mad at you  
 23 (0.7)  
 24 Lei: No I don't miss [that  
 25 Pau: [.hhuh

Extract 11 represents a lengthy and problematic portion of the session, with Paula resorting to reporting her assumptions and experiences (13-17) to account for having asked Leif whether he misses drinking. This troubled interaction is brought on by Leif's quick, minimising response to Paula's therapeutic question in line 1. Paula produces six turns during the course of Extract 11b before Leif responds vocally, indicating strong disaffiliation by Leif with Paula's assertions. Paula's turn at line 11 ('Really?') challenges Leif's 'no' response (lines 7-9), and he shows no uptake, even during a 3.8 second gap

(line 21), until Paula attempts humour at line 22. At this point, Leif responds favourably and in agreement (line 24).

*[4.1.2] Quick Responses that Challenge*

A quick response is also a feature of client responsive actions that challenge the assumptions of the therapist's initiating action. Extract 12 follows directly from Extract 11 and contains three instances of Leif challenging Paula's assumptions (lines 3, 11, and 29). Once Paula has received confirmation from Leif about something that he does not miss (Extract 2b, line 24), she adds to the list of things that she asserts that Leif does not miss 'you don't miss (1.1) falling down because of drinking'. Leif comes in immediately with 'well I didn't usually do that' (line 3). This is significant, because, as I have noted, micropauses and inter-turn silences tend to be approximately 0.4 seconds between Leif and Paula. Leif's immediate response is thus very quick in this context. Paula produces an open class repair initiator (Drew, 1997) in line 4, which is associated with inapposite or disaffiliative prior turns (i.e. Leif's disagreement in line 3). Leif specifies the referent of 'do that' as 'fall down' in line 5, thus providing a repair solution to Paula's open class repair initiator. Paula now must deal with a declined interpretation. She has a number of options to regain the alignment and affiliation demonstrated in Extract 11b, lines 22-25. The one she chooses is to reformulate her perspective so as to appear as though she meant something in particular with which Leif can agree all along, that his falling down was not because he was drunk but because he was in withdrawal (line 7), showing hesitation and reformulation that invokes Leif's doctor's authority (lines 8-10) (cf. Ellis, Kitzinger, & Berger, under review). Leif refuses this interpretation with 'that was really' in terminal overlap with Paula (line 11). This turn challenges Paula's assertion and comes even more quickly than his previous challenge. Paula offers a collaborative completion (line 12), reasserting her version of events 'that was kind of a time limited thing', effectively



interrupting Leif's TCU as this too is refused. Leif continues his turn in overlap with Paula 'that was really ataxia' (line 13). After a 2.2 second gap (line 14), Leif produces an increment 'so it was kind of a clugeon<sup>21</sup> of ((gesture))' (lines 15-17). Paula produces 'yeah' at line 18, claiming affiliation and alignment with Leif's version of events.

Extract 12a [PL 101609]

- 1 Pau: You don't miss: (1.1) falling down becau::se of  
 2 drinking  
 3 Lei: W' I didn't usually do that  
 4 Pau: Pardon  
 5 Lei: I didn't usually fall down  
 6 (0.8)  
 7 Pau: Even with the PAWS<sup>22</sup>  
 8 (.)  
 9 Pau: Even when um (0.8) your doctor was talking  
 10 about the post withdrawa[l  
 11 Lei: [That that really .hh ['s  
 12 Pau: [That kind of was a time limi[ted thing  
 13 Lei: [That was really ataxia  
 14 (2.2)  
 15 Lei: An so it was kind of a \*clugeon\* of ((\*brings  
 16 hands together interlocking fingers slightly  
 17 during this word))  
 18 Pau: ↓Yeah

Leif continues his prior turn after a 1.2 second gap (lines 20-22) with 'two things but I y'know I didn't fall down yeah'. After a 3.2 second gap, Leif reformulates his turn at line 22 with 'yeah I didn't (2.3) you know puke all over myself or anything like that uh' (lines

<sup>21</sup> As noted in Chapter 3, 'clugeon' is a spoken-only regional word meaning approximately a tight jumble of something. It has been transcribed to resemble 'bludgeon', with which it rhymes.

<sup>22</sup> Post-alcohol withdrawal syndrome

24-25. After a brief gap, Paula says ‘well no you don’t miss the disgusting things I guess at all but’. This turn claims both affiliation and alignment but demonstrates instead affiliation with something Leif did not say and misalignment, a lack of intersubjectivity with Leif’s course of action (i.e., Leif has said this didn’t happen, and Paula has implied that it did and that he would not miss these ‘disgusting things’). Leif rejects Paula’s interpretation ‘y’I didn’t- really didn’t-’

Extract 12b [PL 101609]

- 19 (1.2)
- 20 Lei: .h two things but [I I=
- 21 Pau: [.hhhhh
- 22 Lei: =>Yknow<I didn’t fall down yeah
- 23 (3.2)
- 24 Lei: Yeah I didn’t (2.3) you know puke all over myself
- 25 or anything like that uh
- 26 (0.9)
- 27 Pau: h well no you don’t miss the disgusting things I
- 28 guess °at all° but w-
- 29 Lei: Y’I didn’t- really di[dn’t-

#### *[4.3] Re-establishing Alignment and Affiliation*

Revisiting Extract 10, reprinted here as Extract 13 with the final line of Extract 12 included, one can identify striking differences between the sequential environment of Extracts 11 and 12 and Extract 10/13. In line 1, Leif is denying Paula’s assertion that he ‘doesn’t miss the disgusting things’ that he said he did not experience. Paula again interrupts him, this time, however, with a therapeutic question ‘what about it was so seductive’ (line 2), which does not presume any particular experiences other than that of being drawn to drinking. There is an immediate, observable switch in the interaction, and Leif performs sincerity in large part through his silence. Line 2 is Paula’s therapeutic question following misaligned

talk about Leif's experiences with alcohol. This is followed by a brief silence at line 3 and the beginning of Leif's answer at line 4, which after many silences and much expansion is fully realised in lines 19-20. Although Leif's answer could continue further, it is at this point that substantial information has been conveyed and at which Paula explicitly affiliates with his response (line 22, 'yeah').

Extract 13 [PL 101609]

1 Lei: Y'I didn't- really di[dn't-  
2 Pau:>> [What about it was so seductive  
3 (0.9)  
4 Lei:>> .hhhh Wl I've been thinking about that,  
5 (0.8)  
6 Lei: ↑I really have  
7 (1.6)  
8 Lei: Um  
9 (1.8)  
10 Lei: MOSTly what I didn- (3.3) mostly what I don't miss  
11 (2.1) is disappointing myself  
12 (0.7)  
13 Pau: Hm  
14 (2.0)  
15 Pau: Hm  
16 (1.0)  
17 Lei: In other words (0.5) spending ti::me (6.6)=  
18 Pau: hmhehhu[h ((cough))  
19 Lei:>> =[basically getting wa:sted (2.3) when I probably  
20 could have been doing s(h)omething else  
21 (.)  
22 Pau: °Yeah°

A similar situation occurs in Extract 14, prior to which Leif and Paula have been discussing what Leif is experiencing as depression. In line 1, Paula challenges Leif's conceptualisation of sleeping in and not getting things done as depression, 'and how do you know it's depression'. Leif immediately responds (line 2) and is unable to complete his turn before Paula produces a self repair (line 3) that is not necessary for understanding the question. Instead, this self repair functions as a display of accountability for having asked the question. By producing it shortly after Leif begins his response 'I don-', Paula is displaying an orientation to Leif's response as dispreferred.

Extract 14 [PL 100109]

- 01 Pau: And how do you know it's depressh: depressed  
 02 Lei:>> I don:-  
 03 Pau: W'I [I mean  
 04 Lei:>> [I  
 05 Pau: Instead of pleasant  
 06 (1.6)  
 07 Pau: What's the quality of it  
 08 (.)  
 09 Lei:>> .I do(h)n't feel good about it  
 10 (2.4)  
 11 Pau: What are you thinking  
 12 (2.2)  
 13 Lei: I'm  
 14 Pau: =( ) don't feel good about it sounds like a thought.  
 15 Lei: I'm I'm pretty much (0.9) feeling like (2.5) you know there's  
 16 really no place for me here (0.5) you know I['m,

In line 4, Leif tries again 'I' but is in overlap with Paula and drops out. Paula ultimately re-issues her question as 'what's the quality of it' (line 7), to which Leif

produces a complete turn ‘I don’t feel good about it’. This turn is likely that which he had previously attempted, because it is prosodically similar and contains disturbances during both words that were present in his earlier partial turns. Disturbances such as aspiration can indicate trouble with a word (Potter & Hepburn, 2010), and in this case indexes prior interactional trouble rather than a problem with the words themselves as part of the turn. In line 11, Paula produces a follow-up question ‘what are you thinking’, indicating that Leif’s response is inadequate. After a 2.2 second gap (line 12), both parties begin speaking almost simultaneously. Although Leif begins first (line 13), Paula’s turn is latched to his first syllable (line 14) and explicitly targets Leif’s response as inadequate ‘( ) don’t feel good about it sounds like a thought’. Although Leif’s second try in responding to Paula’s last question comes immediately following her turn at line 14, he pauses multiple times, and the focused talk he produces in lines 15-16 continues beyond this extract (data not shown). By engaging in silence, Leif makes moves toward realignment with Paula.

#### *[4.4] Increments and Reformulated Turns versus Continuation*

The therapist generally provides ample time for responding but does, on occasion, produce increments or redoes a turn. This is particularly evident in Extract 11, partially reproduced here as Extract 15.

Extract 15 [PL 101609]

- 1 Pau: hh do you miss it
- 2 ((lines deleted, repair sequence))
- 3 (0.5)
- 4 Lei: No
- 5 (0.4)
- 6 Lei: Not really
- 7 (0.7)
- 8 Pau: Really?

9 (1.1)

10 Pau: Cause I thought for most people it's sorta like

11 (0.9) s:moking (.) you miss tha:t

12 (0.8)

13 Pau: There are moments when you just (0.8) boy it'd be

14 a perfect moment for a cigarette you know

15 (0.9)

16 Pau: You don't miss: (1.9) °you don't miss drinking°

17 Pau: You don't miss time went on

18 (3.8)

19 Pau: You don't miss Patty getting mad at you

In Extract 15, following a repair sequence (line 2), Leif responds 'no' (line 4) to Paula's yes/no interrogative 'do you miss it [drinking]' (line 1). He expands this answer in line 6 as 'not really', which mitigates the dispreferred response. Although missing drinking is negatively valued, in the context of psychotherapy, a simple 'no' to a therapeutically relevant question blocks the therapist's therapeutic agenda. After a brief silence (line 7), Paula expands this sequence with 'really' with rising intonation, demonstrating surprise and an element of disbelief. Leif does not respond to Paula's turn, and a 1.1 second gap develops (line 9). Paula takes additional turns at lines 10-11, 13-14, and 16-17, whilst Leif still does not respond during moderate silences of 0.8-1.9 seconds. At line 18, a 3.8 second silence occurs before Paula takes another turn. Paula's turns claim a degree of continuity between each successive one. Compare this to Extract 12b (partially reproduced here as Extract 16), in which Leif breaks the silences rather than Paula.

Extract 16 [PL 101609]

1 Lei: =>Yknow<I didn't fall down yeah

2 (3.2)

3 Lei: Yeah I didn't (2.3) you know puke all over myself

4                    or anything like that uh

Following Leif's turn at line 1, which is neither conditionally relevant nor incomplete, a 3.2 second silence ensues (line 2). Leif then self-selects at lines 3-4 with a continuation of line 1, using the same first word in line 3 as his last word in line 1 and on the same topic. In both Extracts 15 and 16, Paula and Leif have employed the same practices of self-selection following a gap. This is a very different practice from that of Leif's use of silence to perform sincerity in Extract 10, reproduced here in full.

Extract 10 [PL 101609]

1    Pau:        [What about it was so seductive  
2                    (0.9)  
3    Lei:        .hhhh Wl I've been thinking about that,  
4                    (0.8)  
5    Lei:        ↑I really have  
6                    (1.6)  
7    Lei:        Um  
8                    (1.8)  
9    Lei:        MOSTly what I didn- (3.3) mostly what I don't miss  
10                    (2.1) is disappointing myself  
11                    (0.7)  
12   Pau:        Hm  
13                    (2.0)  
14   Pau:        Hm  
15                    (1.0)  
16   Lei:        In other words (0.5) spending ti::me (6.6)=  
17   Pau:        hmhehhu[h ((cough))  
18   Lei:                =[basically getting wa:sted (2.3) when I probably  
19                    could have been doing s(h)omething else  
20                    (.)

21 Pau: °Yeah°

In Extract 10, Paula asks Leif a yes/no interrogative (line 1), which is conditionally relevant. Leif's delay at line 2, rather than indicating a dispreferred response as in ordinary conversation, displays consideration of the question and participating in the therapeutic process. He produces a preliminary to responding 'well I've been thinking about that ... I really have' (lines 3-5). Following this action that is composed of two TCUs, a 1.6 second silence develops. This is only broken by Leif's 'um' at line 7 and is followed by another 1.8 seconds of silence. Leif has not yet produced his actual response to Paula's question at line 1. However, he has shown orientation to it and consideration of it. He begins his actual response at line 9 and initiates self-repair mid-TCU. The 3.3 second silence in line 9 is allowed to develop by both parties, as is Leif's next pause in line 10 of 2.1 seconds.

Leif has the right to continue his turn and the obligation to provide an adequate response to Paula's question. Leif's 'mostly what I don't miss is disappointing myself' does not fully answer Paula's question; it lacks the specificity of a fully developed idea in the therapeutic process. Paula, however, at this point acknowledges Leif's response as a fully fledged response but not as one that is ready for a therapeutic intervention such as an interpretation. Instead, she also shows that she is considering – not her own response, but Leif's reported thoughts on the matter. Paula says 'hm' twice during an otherwise silent approximately 4 seconds (lines 11-15). Leif then offers further specification at line 16 'in other words spending time', which could be a grammatically and syntactically complete turn but would be more complete with a gerund phrase (which is then added as '... getting wasted ...'). The 6.6 second silence in line 16 is thus transcribed as Leif's pause. It is also treated as such by Paula, who merely coughs (line 17). Leif continues his turn in lines 18-19 in overlap with Paula's cough 'basically getting wasted (2.3) when I probably could have been doing something else'. At this point, Paula produces 'yeah' (line 21), which



displays both alignment and affiliation with Leif's response. It is a sequence closing third that does not make closure mandatory but instead can serve multiple functions that participants can exploit as they continue to progress the conversation. By carefully balancing her own talk and silence, Paula is able to encourage Leif to talk and relinquishes some control of the interaction, consistent with humanistic approaches to therapy. By engaging in this balancing, Paula also maintains a contemplative demeanour that is consistent with her approach to spiritually oriented psychotherapy.

### **[5] Conclusion**

From a structural point of view, the silences that are attributable to Leif in Extract 1 occur as a consequence of his not yet producing the implicated response but orienting to its requirement nonetheless. The silences are also simultaneously responsible in part for conveying the gravity of Leif's response. Just as silences that are associated with dispreferred responses in ordinary conversation demonstrate consideration and perform reluctance, so do silences associated with preferred responses to therapeutic interventions. A prompt response in ordinary conversation is associated with preferred responses, along with prompt dispreferred responses being accountable and/or rude. Additionally in ordinary conversation, immediacy and spontaneity are valued in the immediate sequential environment. When a client responds to an intervention such as a therapeutic question quickly, however, the therapist often treats this as inappropriate to the sequential environment and is likely to challenge the client's response.

The current study is one of the few works that look empirically at silence in a psychotherapeutic context. Previous work in this area has been largely conjectural and based on the memories of the practitioner, and many have described silence as problematic and representative of mental illness and resistance. By approaching silence that occurs in

actual, recorded interactions and from a specialist technical perspective such as conversation analysis, we can move forward from the top-down approach of psychoanalytic theory that dominates the literature on silence in psychotherapy at present. By recognising that silence is not necessarily problematic in psychotherapy, and especially that silence may be a sign that the client is in fact giving due consideration to a therapeutic intervention rather than 'resisting' it, psychotherapists and members of related professions such as counselling and psychiatry can achieve better understanding of and rapport with clients. In this chapter, I have examined a set of interactions that are structured through talk yet have special ways of dealing with silence compared to ordinary conversation. Interactions in which silence is treated as unproblematic may constitute a special kind of interaction. In the next chapter, this proposal is explored and critiqued.

## Chapter 5

### Do 'Continuing States of Incipient Talk' Exist?<sup>23</sup>

---

<sup>23</sup> The version presented in this chapter was written by me, and the quantitative design is my own. Rowena Viney has been involved in discussions, given feedback on the written work, and second-coded the data. John Rae has been involved in discussions and given feedback on the written work. The version that has been submitted for publication includes both as co-authors. It has also been presented as:  
Berger, I., Viney, R., & Rae, J. (2012, March 22). *Toward a cohesive conceptualisation of incipient talk*. Discourse-Communication-Conversation. Loughborough, UK.

### **Abstract**

Reference to the concept of 'incipient talk', or more often 'continuing states of incipient talk', is centred around a brief and speculative note in the conclusion of Schegloff and Sacks's 1973 paper on closings in conversation. Their note contains no data or analyses, yet it has been taken as a set of facts by numerous researchers. Whilst terms may be used initially in a less technical sense, they can take on an authority of their own over time.

Many authors do not even cite Schegloff and Sacks's use of the phrase. This chapter aims to challenge the assumptions made in the adoption of 'incipient talk' or 'continuing states of incipient talk' by showing the range of ways in which these terms are used within conversation analysis and related fields of enquiry and by making explicit the contradictions of stated and/or implied reasoning within and across studies that use these terms. By approaching the issue with a focus on the ways in which the terms are used in the literature, one can see the degree to which idiosyncratic uses and definitions dominate the field.

We present results from a content analysis of 72 papers that use the phrase 'incipient talk' and show that multiple, disparate usages and definitions exist. We then discuss some of the uses and compare 'incipient talk' to possibly related concepts 'open state of talk', 'unfocused interaction', and 'islands of talk'/'Gesprächsinselfn'. We provide suggestions for future research in clarifying whether 'incipient talk' exists and what it would be.

**[1] Introduction**

One explanation for the unproblematic nature of silence in social interaction is that the interaction constitutes a 'continuing state of incipient talk'. Discussion of the concept of the 'continuing state of incipient talk', or sometimes simply 'incipient talk', is centred around a brief, speculative note in the conclusion of Schegloff & Sacks's (1973) paper on closings in conversation:

... What we are really dealing with is the problem of closing a conversation that ends a state of talk. It does not hold for members of a household in their living room, employees who share an office, passengers together in an automobile, etc., that is, persons who could be said to be in a 'continuing state of incipient talk'. In such circumstances, there can be lapses of the operation of what we earlier called the basic features; for example, there can be silence after a speaker's utterance which is neither an attributable silence nor a termination, which is seen as neither the suspension nor the violation of the basic features [(1) at least, and no more than, one party speaks at a time in a single conversation; and (2) speaker change recurs (p. 293)]. These are adjournments, and seem to be done in a manner different from closings. Persons in such a continuing state of incipient talk need not begin new segments of conversation with exchanges of greetings, and need not close segments with closing sections and terminal exchanges. ...

Their note contains no data or analyses, nor does Sacks (1992) ever address this issue in his lectures from the same period. Yet the concept of a continuing state of incipient talk has been taken as an empirical finding by numerous researchers. Neither 'incipient talk' nor 'continuing states of incipient talk' have ever been defined through the systematic, large scale analyses that have been at the heart of conversation analytic research since its

inception. Individual authors have nonetheless idiosyncratically defined incipient talk and used these definitions to support their analyses. Whilst terms may be used initially in a less technical sense, they can take on an authority of their own over time. For example, Macbeth (2004) critiques the usage of 'repair' to refer to pedagogical corrections in classrooms. This chapter aims to challenge the assumptions made in the adoption of 'incipient talk' or 'continuing states of incipient talk' by showing the range of ways in which these terms are used within conversation analysis and related fields of enquiry and by making explicit the contradictions of stated and/or implied reasoning within and across studies that use these terms. By approaching the issue with a focus on the ways in which the terms are used in the literature, one can see the degree to which idiosyncratic uses and definitions dominate the field. We begin with a quantitative content analysis in order to explore the complex inter-linkages of how these terms are used by authors and how this may vary across publication year and type. We then discuss specific themes in how authors define and use 'incipient talk', compare potentially related concepts in social research, and suggest future research directions.

## **[2] Content Analysis**

To assess the scope and range of usage of 'incipient talk', we performed a content analysis on interactional features that are described as incipient talk and definitions (if any) for incipient talk that are given in these publications. We predicted that multiple clusters would emerge, indicating quantifiable variations in how 'incipient talk' is used and defined and that publication type would have no effect on these clusters. We predicted also that more recent publications would have membership in more clusters.

Because 'incipient talk' is typically mentioned in passing or as minor analytic points, abstract databases are not appropriate for this study. For example, PsycINFO gave only

two results for “incipient talk”, whereas Scopus gave four. Furthermore, the paper from which ‘incipient talk’ originates is itself not included in major databases. Therefore, we used Google Scholar, which has full text capability and searches a range of sources. Because we are not using impact factors within the search results in our analysis, Google Scholar does not present significant methodological challenges. Content analytic methodology precludes the addition of entries not found in a particular search. To reflect what people find when they search, we have not excluded any document types a priori.

A search on Google Scholar for “incipient talk” without any advanced search options on 18 November 2011 gave 89 results. Although literature might exist that uses ‘incipient talk’ in its abstract, text, or title that were not included in the search results, content analytic methodology does not allow for selectively adding cases that are subsequently located outside the search parameters. Of these results, eight entries were reprints or duplicate entries and were excluded. Where a conference paper and a journal article or book chapter appeared, the conference paper was treated as the duplicate entry and excluded except where there were substantial differences in the overall content. We excluded two entries that did not discuss incipient talk but were included in the search results simply because they cited a paper with ‘incipient talk’ in the title. Two entries were excluded, because the sources could not be located. One undergraduate dissertation was excluded due to being of particularly low quality and because we were unable to gain access to the full document. Two results pointed to index entries for other results and were excluded.

These exclusions resulted in the final sample containing 72 entries (Appendix J). This included 37 journal articles, 15 book chapters, six books, three conference proceedings, five postgraduate theses, four commentaries in journals, one unpublished

manuscript, and one research proposal. Publication years ranged from 1973 to 2011 with a median year of 2006, indicating a recent surge in interest in 'incipient talk'. One entry, the research proposal, had no date. Most authors had one to three entries. However, the most prolific author was Emanuel A. Schegloff with 10 entries, and a group of frequent co-authors were listed on seven, seven, and five entries each. Entries were predominately in English. Of the included entries, 66 were in English, three in French, and one each in Portuguese, Dutch, and Swedish.

### *[2.1] Variables*

Nineteen variables were generated through examination of all 72 included entries. All 72 included entries were then coded for these 19 variables which were related to the explicit or implicit usage of the term 'incipient talk'. Variables were generated through examination of all definitions and uses of 'incipient talk' in the sample. All entries that used 'incipient talk' to refer literally to talk that was beginning were negative for all other descriptive variables. Excluding this variable from the analysis left 18 variables.

Publications were also coded for eight variables related to the descriptions themselves (e.g., who is cited). Tables 1 and 2 shows these variables and how frequently they occurred.



Table 1. Properties of Descriptions	
Variable	n (%)
Cites Schegloff & Sacks (1973)	40 (55.56%)*
Correctly cites Schegloff & Sacks (1973)	34 (47.89%)* / (85.00%)**
Mis-cites Schegloff & Sacks (1973)	6 (8.33%)* / (15.00%)**
Cites other paper(s)	19 (26.39%)
Defines 'incipient talk'	24 (33.33%)
'Incipient talk' is taken for granted	41 (56.94%)
Acknowledges lack of evidence	1 (1.38%)
Discussed anomalies otherwise explained by CA	9 (12.50%)
Contradicts self	2 (2.78%)
High sensitivity to silence (less than 5 sec.)	4 (5.56%)
* N = 71 due to exclusion of Schegloff & Sacks (1973)	
** N = 40, total papers that cited Schegloff & Sacks (1973)	

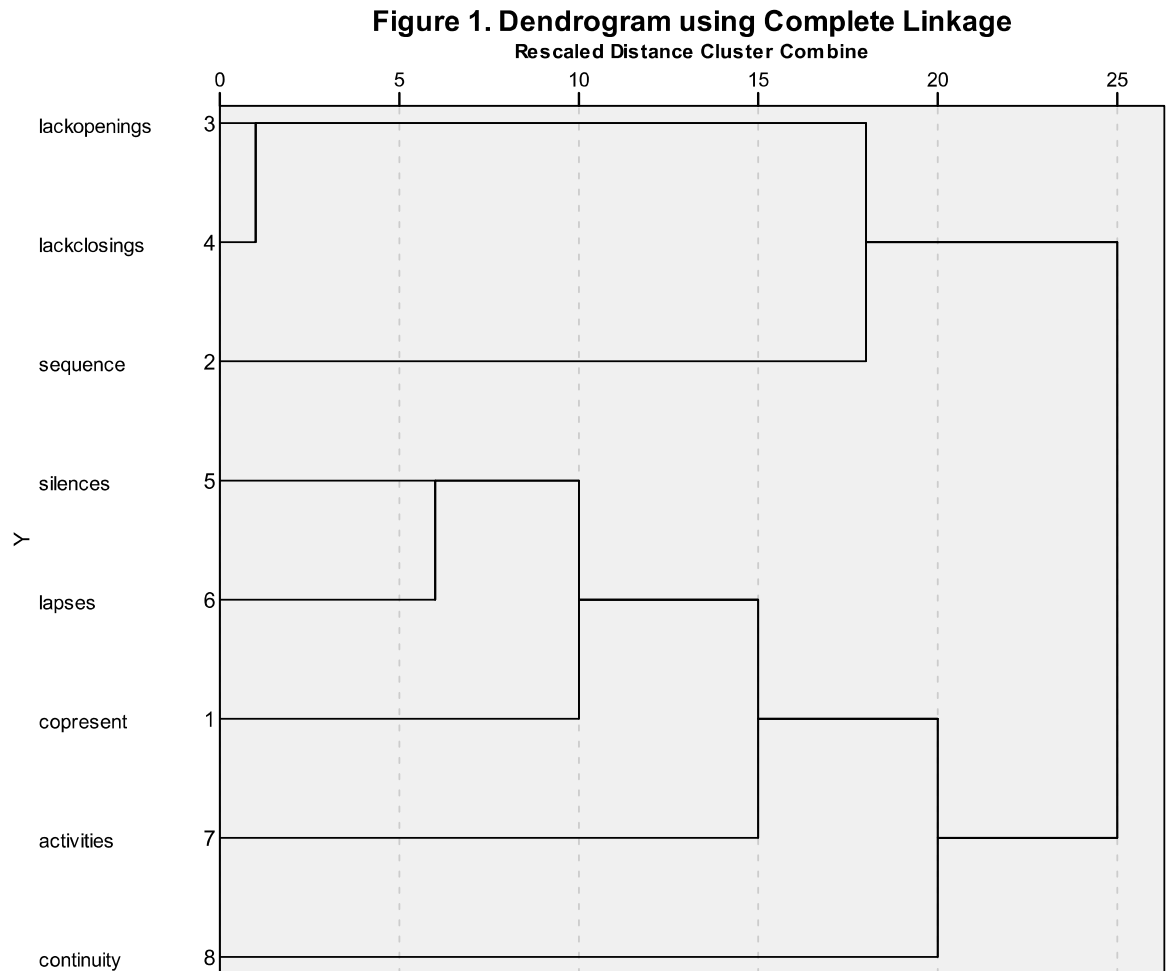
Table 2. Descriptive Variables	
Variable	n (%)
Co-present	25 (34.72%)
Lacks openings	22 (30.56%)
Lacks closings	20 (27.78%)
Involves silences generally	23 (31.94%)
Involves lapses	33 (45.83%)
Sequence organisation applies less	6 (8.33%)
Involves gesture	1 (1.38%)
Doing other activities/talk is not central	19 (36.39%)
Multi-party engagement	3 (4.17%)
Summons-answer or attention-getting	3 (4.17%)
Silence is different from gaps and pauses	3 (4.17%)
Continuity across lapses	7 (9.72%)
Lacks continuity across lapses	2 (2.78%)
Unacceptable or contrasted with acceptable interaction	1 (1.38%)
Participants in very close proximity	1 (1.38%)
'Conversational' as opposed to 'transactional'	1 (1.38%)
No clear beginning or end	2 (2.78%)
Set temporal beginning and end	1 (1.38%)

### [2.2] Results

Twenty (27.78%) entries were randomly selected for inter-rater reliability analysis, which showed overall excellent agreement for the variables that were subsequently included in the cluster analysis (copresent  $\kappa = .600$ , lapses  $\kappa = .667$ , sequence  $\kappa = .773$ , silences  $\kappa = .829$ , lackopenings  $\kappa = .857$ , lackclosings  $\kappa = .828$ , activities  $\kappa = .886$ ).  $\kappa$  could not be

calculated for continuity in the sub-sample due to not being present in those 20 papers. In order to improve agreement further, the entire dataset was second coded by the same coder and disagreements resolved through discussion. Following complete dataset second coding, ten entries were found to be too vague in their descriptions to qualify for any of the codes. A hierarchical cluster analysis on the codes and using complete linkage (furthest neighbour) clustering and Dice measurement was performed due to the binary nature of the data. Variables that were present in three or fewer entries were excluded, because they can create arbitrary clusters by introducing inherently equal distances into the proximity matrix. Only eight of the uses or definitions present in the data could be included in the cluster analysis and were represented in five or more papers. Variables with three or fewer positive responses were excluded from the cluster analysis in addition to the 'talk that is beginning' variable, which did not share positive responses to any other codes.

Cluster 1 contained *lacks openings, lacks closings, and sequence organisation applies less*. Cluster 2 contained *silences (not specifically lapses) are involved, lapses are involved, co-present, doing other activities/talk is not central, and continuity across lapses*. Figure 1 shows the dendrogram for the cluster analysis. Clusters were identified at a rescaled distance of 20, a conservative cut-off given the large distances before *sequence organisation applies less, doing other activities/talk is not central, and continuity across lapses* were absorbed into the respective clusters.



Scores were calculated for these clusters on the basis of number of positive responses divided by the number of variables in the cluster to give scores between 0 and 1 (Cluster 1  $M = .22$ ,  $SD = .32$ ; Cluster 2  $M = .30$ ,  $SD = .28$ ). Both clusters' scores were non-normally distributed (Cluster 1 Shapiro-Wilk = .70,  $p < .001$ ,  $Z_{\text{skew}} = 3.92$ ,  $Z_{\text{kurtosis}} = -.25$ ; Cluster 2 Shapiro-Wilk = .88,  $p < .001$ ,  $Z_{\text{skew}} = 2.48$ ,  $Z_{\text{kurtosis}} = -.73$ ). Cases were identified whether they fell into each cluster (i.e. whether they scored higher than zero), and membership in clusters was non-exclusive. Cluster 1 accounted for 27 (37.50%) cases, and Cluster 2 accounted for 50 (69.44%) cases. Twenty one cases (29.17%) did not fall into either cluster, including the ten that could not be adequately coded and the nine that only used 'incipient talk' to literally mean talk that was beginning. Twenty five (32.72%) cases fell into one cluster, and 26 (36.11%) fell into both clusters.

Twenty non-continuous years ranging from 1973 to 2009 were represented in the data set. Years were not normally distributed (Shapiro-Wilk = .81,  $p < .001$ ;  $Z_{\text{skew}} = -7.79$ ,  $Z_{\text{kurtosis}} = 13.86$ ). The number of clusters per year and publication year were not significantly correlated ( $r_s = -.21$ ,  $p = 0.08$ ). A Kruskal-Wallis test was used to evaluate whether differences in Cluster 1 scores, Cluster 2 scores, and number of memberships in clusters existed across publication type, which were non-significant ( $p = .82$ ,  $p = .84$ , and  $p = .86$ , respectively).

Due to over half of the definitions or usages of 'incipient talk' not occurring in five or more cases and thus not being eligible for the cluster analysis, we computed a total score of the 18 variables (not including literal usage). Total scores are a measure of elaboration that entries have in their usage or definition of incipient talk/continuing states of incipient talk; those that have more definitions/usages are considered more elaborate. However, this does not necessarily represent consistency with Sacks & Schegloff (1973) nor quality of usages and definitions. Elaboration scores were not normally distributed (Shapiro-Wilk = .91,  $p < .001$ ;  $Z_{\text{skew}} = 1.69$ ,  $Z_{\text{kurtosis}} = -1.13$ ) and ranged from 0 (those that only used the term literally) to 8. Elaboration score, like clusters, was not significantly correlated with year, nor did it differ by publication type (Kruskal-Wallis  $p = .89$ ). Mann-Whitney U tests were used for all entries and usages/definitions for which  $n \geq 4$ . Elaboration was associated with a number of features of the entries (Table 3), particular definitions and usages (Table 4), and cluster membership.

Table 3. Differences in Total Definitions/Usages for Properties of Descriptions			
Variable	Yes M(SD)	No M(SD)	p
Cites Schegloff & Sacks (1973)	3.40(1.84)	1.03(1.43)	< .001
Correctly cites Schegloff & Sacks (1973)	3.50(1.81)	1.32(1.65)	< .001
Mis-cites Schegloff & Sacks (1973)	2.83(2.04)	2.32(2.05)	.41
Cites other paper(s)	3.21(2.18)	2.11(1.94)	.06
Defines 'incipient talk'	3.79(2.15)	1.71(1.61)	< .001
'Incipient talk' is taken for granted	2.34(1.94)	2.48(2.20)	.89
Discussed anomalies otherwise explained by CA	2.67(2.18)	2.37(2.04)	.80
High sensitivity to silence (less than 5 sec.)	3.25(1.89)	2.35(2.06)	.36

Table 4. Differences in Total Definitions/Usages for Individual Definitions/Usages			
Variable	Yes M(SD)	No M(SD)	p
Co-present	4.24(1.56)	1.43(1.54)	< .001
Lacks openings	4.18(1.62)	1.62(1.70)	< .001
Lacks closings	4.60(1.27)	1.56(1.61)	< .001
Involves silences generally	4.22(1.73)	1.55(1.58)	< .001
Involves lapses	3.91(1.68)	1.13(1.34)	< .001
Sequence organisation applies less	4.83(2.04)	2.18(1.91)	.008
Doing other activities/talk is not central	3.68(1.95)	1.94(1.90)	.002
Continuity across lapses	4.14(1.57)	2.22(2.01)	.02

Citing Sacks and Schegloff (1973) at all was associated with higher elaboration, as was correctly citing the paper (compared to the entire sample). Scores did not differ, however, for those who mis-cited the paper and those who did not in the entire sample.

Unsurprisingly, papers that defined incipient talk scored higher on elaboration than those who did not. All eight of the variables used in the cluster analysis were associated with higher elaboration, as was cluster membership. Entries that were not members of either cluster ( $M = .10$ ,  $SD = .30$ ) were significantly lower in elaboration than those that belonged to at least one cluster ( $M = 2.25$ ,  $SD = 1.66$ ). Additionally, both Cluster 1 ( $r_s = .68$ ,  $p < .001$ ) and Cluster 2 ( $r_s = .88$ ,  $p < .001$ ) scores were strongly correlated with elaboration scores.

### [2.3] Discussion

Even a casual glance at Tables 1 and 2 reveals how varied authors' use of 'incipient talk' is. Closer inspection reveals contradictions as well. Two entries (2.78%) were self-contradictory in their definitions of 'incipient talk', and six entries (8.33% of total; 15.00% of those who cited *Opening Up Closings*) mis-attributed claims to or mis-quoted Schegloff and Sacks (1973). Seven (9.72%) entries claimed that 'incipient talk' involves continuity across lapses, whilst two (2.78%) claimed that it lacks continuity across lapses. Moreover, 21 (29.17%) of publications did not fall into any of the clusters of commonly used definitions, and over half of the variables ( $N = 10$ , 55.56%) had to be excluded from the cluster analysis due to too few publications using them. Despite these obvious flaws in the construct of continuing states of incipient talk, only one entry (Stivers & Rossano, 2010a) acknowledged that there is a lack of evidence regarding categorising interactions as 'incipient talk'.

Two of the three hypotheses of this content analysis were supported. We had predicted that multiple clusters would emerge, and two distinct clusters were found using a conservative methodology. We predicted that publication type would not have an effect on the cluster scores, and it did not. We predicted that more recent publications would be

members of more clusters, but this was not the case. Due to the large proportion of variables that did not qualify for cluster analysis, we conducted an analysis of elaboration using the total number of definitions/usages present in each entry. Unsurprisingly, entries that defined incipient talk explicitly had higher elaboration scores than those that did not. Entries that cited other papers, particularly Sacks and Schegloff (1973), were also higher in elaboration. Every definition/usage that was eligible for the cluster analysis was associated with higher levels of elaboration. Moreover, entries that were members of at least one cluster had much higher elaboration scores than those that did not belong to either cluster. This is potentially a positive sign in that although there is a wide range of definitions and usages as well as low membership in clusters (nearly one-third fall into neither cluster), those that are most elaborate in how they use and define incipient talk are using more similar concepts than those who are not. This is not to say, however, that there is consensus within clusters, as membership in clusters was defined by using at least one of the concepts that fell within each cluster and the mean cluster scores were quite low (.22 and .30 respectively). Individual authors are not marrying these fragmented and divergent uses into a complex theoretical construct, but rather the field remains fragmented in how it uses 'incipient talk'. This leads us to address 'incipient talk' through a systematic review that will encompass the remainder of this chapter.

### **[3] Silence and Sequences**

As almost all of the uses of 'incipient talk' involve patterns of silence in some form, it is necessary to review the kinds of silence and their positions in conversation. These definitions follow a typology that is consistent with conversation analysis. The terms we will use to refer to silence are *pauses*, *gaps*, and *lapses* (e.g., Sacks, Schegloff, and Jefferson, 1974). A *sequence* is a unit of talk that has at least a first pair part (e.g., a question) and a second pair part (e.g., an answer) (e.g., Schegloff, 1968; Schegloff & Sacks,



1973). They are made up of turns and may be expanded before, between, or after these base pairs (see also Schegloff, 2007). Silence can occur at any point in a sequence or outside of it, and whether a given silence is called a pause, a gap, or a lapse depends on where it occurs. In this section, we use data extracts to illustrate different kinds of silences. The data presented were collected with the participants' informed consent and have been pseudonymised for publication.

### *[3.1] Pauses*

Pauses occur within turn constructional units (TCUs). These silences effectively belong to the speaker within whose turn it occurs. They have the right to continue speaking and may be accountable for not doing so.

Extract 1 [PL 121809]

01 Lei:>> So far I've managed to:: uh (2.7) keep Patty from  
02           going absolute totally bonkers paranoid,

### *[3.2] Gaps*

Gaps occur within sequences, outside of turns. A delay in responding is often taken to indicate that a dispreferred second pair part might be forthcoming. Recipients to whom a first pair part is addressed have the right and obligation to respond. In other words, the silence is attributable to a particular party (which may include multiple people).

Extract 2 [CA 062109]

01 Al:     Have you seen or heard the neighbour downstairs  
02         >> (0.7)  
03 Cor:    No I've not

*[3.3] Lapses*

Lapses are silences that occur outside of sequences. They can be attributed to no party, and they may be ended by any speaker. Although the term 'lapse' has connotations of a long silence, lapses can be of any length.

Extract 3 [PL 091609]

01 Pau: =I also kept his e-mail (1.0) with the icons on  
 02 the bottom >and I can't remember: if I shot it  
 03 directly from the icon: or >I think I sav- I I shot  
 04 it from my Word I just did it as a=  
 05 Lei: Yeah  
 06 Pau: =Send[ file ( )  
 07 Lei: [Well either either or because it was coming (1.8)  
 08 that way it probably (0.7) it probably didn't  
 09 (1.9) work  
 10 >> (1.9)  
 11 Pau: So how are ya

As we will show in the subsequent sections, although silence is a common feature of interactions described as 'incipient talk', how silence is involved is widely divergent. Some authors include any silence, including frequent pauses, as incipient talk, whilst others include only lengthy lapses. Similarly, although silence is a common feature, it is not always included in definitions of incipient talk (if a definition is provided). We will extract a definition based on Schegloff & Sacks (1973) before examining features of interactions that have been described as incipient talk and explicit definitions.

**[4] Unpacking Incipient Talk**

Returning to our extract from Schegloff & Sacks's note, we can identify some key features of conversations that could be characterised by the proposed concept of incipient talk.

... members of a household in their living room, employees who share an office, passengers together in an automobile, etc. ... lapses of the operation of what we earlier called the basic features; for example, there can be silence after a speaker's utterance which is neither an attributable silence nor a termination, which is seen as neither the suspension nor the violation of the basic features ... These are adjournments, and seem to be done in a manner different from closings ... need not begin new segments of conversation with exchanges of greetings, and need not close segments with closing sections and terminal exchanges ... (Schegloff & Sacks, 1973)

Simplified, we have extracted three features:

1. Parties are in proximity not by virtue of talking to one another but by other circumstances.
2. The basic features of conversation are not present at all times. The basic features are that (a) at least, and no more than, one party speaks at a time in a single conversation; and (b) speaker change recurs (Schegloff & Sacks, 1973, p. 293). For these to be absent, no party would be speaking, and no change of speaker would be in progress. In other words, lapses (as defined in the previous section) occur.
3. Greetings and closings are not necessary to begin or end spoken interaction.

### **[5] Characterisations of Incipient Talk**

We now turn to some of the uses of 'incipient talk' for an in-depth look at how authors are using them. Although these do not represent all eighteen uses identified, they reflect the great variation in usage that has already been identified quantitatively.

*[5.1] Lack of Openings and/or Closings*

Work on push-to-talk radio use (Woodruff & Aoki, 2004a,b; Szymanski, Aoki, Vinkhuyzen, and Woodruff, 2006) has taken an entirely different stance on the idea that incipient talk is talk where closings do not necessarily end the conversation. They do not treat incipient talk as being defined by lack of closings, but take a softer stance and assert that openings and closings are *rare* and that lapses become relevant at sequence completion. It is not clear how they reason that lapses can have relevance, as lapses are a state of lack of interaction.

A common finding is that the continuous availability of an open channel facilitates a bursty conversation style in which formal conversational openings and closings are infrequent even when long lapses occur between natural sequences of turns at talk ... the open channel puts the participants into what Schegloff and Sacks referred to as a "continuing state of incipient talk", largely obviating the need for telephone-style openings and closings. (Woodruff & Aoki, 2004a, p. 412)

They do, however, define incipient talk, albeit with more interpretation than their source provides. If participants do not use greetings or 'goodbyes' at all during 'continuing states of incipient talk', then one must consider where the boundaries for incipient talk and other talk would lie.

Schegloff & Sacks 1973 described a continuing state of incipient talk as consisting of two elements: a form of overall structural organization of conversation, and an interactional circumstance. Structurally, in a continuing state of incipient talk, turn-by-turn talk occurs, then lapses, then occurs again. Unlike single, bounded face-to-face encounters, speakers in an incipient state of talk do not use greetings to initiate talk, nor do they issue goodbyes prior to a lapse in talk. In a continuing state of incipient talk, once talk has been engaged, further turn-by-turn talk is relevant and may be initiated at any time. The interactional circumstance includes the

participants' copresence and other contingencies for the parties' actions that are independent of the overall structural organization of the participants' talk.

(Szymanski et al., 2006, p. 394)

### *[5.2] Lack of Conditional Relevance*

Szymanski et al. go on (ibid. p. 414) to assert that incipient talk can have delayed or absent responses without invoking consequences, citing Schegloff & Sacks (1973), but relaxed conditional relevance was not one of the features mentioned in Schegloff & Sacks's note on the topic. Their study additionally only addresses push-to-talk, not co-present or telephone conversations. Couper-Kuhlen's (2010) claim that sequence organisation does not apply during 'nonfocused' interactions is consistent with Szymanski et al.'s claim, but is, however, not supported by any data.

### *[5.3] Silences are Common*

Voilmy & Ting (2007, pp. 2, 19) discuss interactions that involve first time wireless internet users. They characterise this environment as a state of incipient talk on the grounds that silences do not close the conversation. However, these silences are typically pauses (attributable silences) rather than lapses and occur where typing or clicking would be interactionally relevant. This is consistent with the structure of sequence and turn-taking organisation more generally, so it is unclear how such an environment constitutes a state of incipient talk. None of the features mentioned in Schegloff & Sacks (1973) are present: Parties must talk to accomplish the tasks at hand (including teaching/learning), one party speaks at a time, speaker change recurs, and parties remain in interaction throughout the session.

Davies (1989), whose research is on Quaker meetings, draws on the observations of others including Goffman (no citation given) that 'speech is not necessarily central to an event, that there are 'other arrangements' in which, for example, people who are on familiar terms with one another may be engaged together in an ongoing activity, and while so engaged occasionally speak aloud.' and defines incipient talk according to Goffman ('open state of talk'): 'What is, however, crucial to the categorization of a MW [meeting for worship] as a continuing state of incipient talk is precisely—as Goffman says—that speech is not central to the event: what is central is silence, or to be more precise, what is claimed to be central is 'celebration', 'baptism' or worship. That is what the MW is for, the silence is the ground on which worship becomes possible and the speech is in some sense heard silence, meditation breaking through, making itself heard.' To credit Goffman (1981) with what Schegloff & Sacks term 'continuing states of incipient talk' (1973) would be retrospective as well as serve to ignore CA's divergence from Goffman's approach (see, for example, Schegloff, 1988). Although Goffman discusses the nature of social engagement extensively throughout his work (e.g., Goffman, 1964), Schegloff and Sacks (1973) do not simply repeat his messages using another vocabulary. As we discuss in the next section, such conflation of concepts is common.

Davies additionally presents a report on a community event, the Quaker Meeting for Worship, which follows a general format of silence except for periodic, spontaneous, *non-addressing, and non-sequentially bound spiritually-motivated utterances*. It is hardly conceivable that every Quaker knows every other Quaker on a familiar enough basis for lengthy silences to occur unproblematically in conversation. The Meeting for Worship is clearly not conversation, although it is unquestionably an instance of social interaction in a culturally meaningful but not conversational way. Davies notes, however, that there are substantial differences between the expectations of a Meeting for Worship and other

situations deemed 'continuing states of incipient talk': 'no piece of ministry begins with greetings, not even the first, and there are no adjacency pairs and therefore again the notion of terminal exchanges does not appear to be relevant.' If applying the criterion of greetings being unnecessary to begin speaking during a continuing state of incipient talk (Schegloff & Sacks, 1973), all interactions *outside of the meeting* between those who attend a given meeting would form an extended interaction that encompasses the meeting as a continuing state of incipient talk.

#### *[5.4] Lapses During Talk-Based Interaction*

As with the push-to-talk study (Woodruff & Aoki, 2004a,b; Szymanski et al., 2006), Raymond (2000) claims more detail than actually present in Schegloff & Sacks (1973) but offers an explicit definition (albeit through mis-quotation of their paper, which appears to be a printing error):

'Schegloff and Sacks (1973) distinguish those occasions when the duration of an interaction is organized solely by what is said within it, as in ordinary telephone conversations, from those occasions where an interaction's duration is shaped by the structure of the event within which it occurs, as between unacquainted seat-mates traveling in an airplane. In the latter, the import of the boundaries of sequences of turns, units through which talk is organized (cf. Schegloff, 1990, 1995), have a different significance for co-interactants than they do in the former. Rather than posing the issue of whether closing the interaction should be initiated – which is not a live option given their forced co-presence – the parties remain in a state of potential mutual engagement. Thus, when the duration of the interaction is structured by something other than the talk itself, a continuing state of

incipient talk 'refers to the lapses that can occur when a sequence is brought to closure and no next party starts' (Schegloff and Sacks, 1973).'

By describing incipient talk as occurring during 'forced co-presence', Raymond neglects situations such as people together in a living room (Schegloff & Sacks, 1973). '... when a sequence is brought to closure and no next party starts' is consistent with lapses more generally than 'incipient talk', so it unclear how this relates to 'incipient talk'. Furthermore, there is no indication of when such a lapse would become no longer an interaction. Could the conversation be resumed without greetings the following week, day, or hour? Would partners saying goodnight and good morning constitute closings and greetings despite 'forced co-presence' of sleeping in the same room?

Raymond also suggests that incipient talk can occur in interactions in which talk is central, as with the live news broadcast:

'The event structure of a live news broadcast recurrently confronts speakers with the potential for such lapses. At the completion of each sequence of actions there may be no further talk by any speaker until either the reporter or the anchor launches into a fresh sequence, reporting on a new development, or updating the audience regarding one that has already taken place. The unfolding character of these broadcasts, then, makes frequent lapses a systematic possibility – a possibility countered only by the pressure to avoid what broadcast journalists refer to as 'dead air', that is, no-one talking. Each potential sequence boundary, each incipient lapse, provides reporters and anchors with a locus for organizing their conduct to display an orientation to the reporter's privileged mandate to authoritatively describe



the scene. That orientation can be displayed through a choice among alternative action types for initiating talk into an incipient lapse.'

This treatment suggests that any lapse automatically creates a state of incipient talk. Robinson (1998), in his work on medical consultations, similarly treats lapses in this manner by using 'continuing state of incipient talk' to refer to a 3.6 second lapse in talk following greetings in which the participants are preparing to start the consultation (i.e. sitting down, etc.) as well as a lapse during which the doctor is preparing to ask the next question in the medical interview. In all of these cases, the lapses are in the service of the ongoing talk-based (and structured) interaction and at a point where the next speaker is reasonably expected to be the doctor/interviewer. 'The doctor and patient again coconstruct a continuing state of incipient talk in which the patient waits for the doctor to initiate a next bit of talk while the doctor prepares to deal with her chief complaint.'

This treatment presents two issues: 1) That doctors and patients co-construct continuing states of incipient talk implies active participation in the development of such a state, which is in Schegloff & Sacks (1973) quite a passive state. 2) In a lapse, no party is accountable for a next action, and any party may end the lapse. However, in Robinson's example, the doctor is writing the patient's response down<sup>24</sup> and has been asking a series of questions which has not been completed. Both Raymond and Robinson's assertions that these silences constitute states of incipient talk are in sharp contrast with environments where parties are involved primarily in non-talk activities and any party can self-select to end a lapse and begin a new sequence, as in Schegloff & Sacks (1973). Moreover, in Robinson's extract, the doctor's activities are accessible to both parties, whereas on the telephone, activities that are necessary to progress the interaction may not necessarily be

---

<sup>24</sup> In the context of a medical interview, it is conceivable that the doctor's writing functions as a sequence closing third, as it is still part of the sequence of asking and getting the question answered. Not writing when one had written answers previously might be grounds for the patient to produce increments.

accessible to both parties. For example, Matthews (2010) describes accounting for silence on customer service lines during activities such as booting a computer (See also Rosenblatt and Li, 2010, on phone conversations with family members). On the telephone, such accounts are necessary to display sustained engagement, whereas in co-present interaction, such preparation for further talk is made visible and lack of speech during these activities does not signify disengagement from the interaction.

#### *[5.5] Multi-Party Engagement*

Jones and Thornborrow (2004) use 'incipient talk' to mean that one party is periodically engaged with various other parties during an event in which there is continuous talk amongst participants. 'There has been a state of incipient talk revolving around the teacher (of the sort discussed above, with queries, helping individual pupils, general reminders).' This is an interesting use of the term, as others have used it to describe situations that involve fewer people and silences among all participants. However, it would seem to fit the three properties that we extracted from Schegloff & Sacks's note. A more thorough analysis of the data in terms of proposed incipient talk would be a worthwhile endeavour.

#### *[5.6] Summons-Answer Sequences*

Some of the studies in the content analysis sample mention the need to get attention from interlocutors in order to begin speaking. Such attention-getting and -checking takes the form of non-vocal practices (e.g., Heath & Luff, 1993). Mondada (2006, section 6.3), on the other hand, states that spoken summons-answer sequences can be used to initiate talk during states of incipient talk. Should such states be actual and classifiable, then this observation could be an important and impactful one for the field. Schegloff & Sacks's description of 'continuing states of incipient talk' implied continuous availability, something to which Mondada's research brings an apparent contradiction. Although talk

during a task is one of the more canonical examples of situations described as a state of 'incipient talk', she has shown that continuous availability cannot be assumed. Schegloff and Sacks consider 'continuing states of incipient talk' to be within a conversation, that is that the conversation has not been terminated and does not require re-opening. If an interlocutor's availability in a conversation is in question in such a way as to require a summons-answer sequence, it is questionable whether the conversation can be considered ongoing. Mondada has provided empirical evidence against the very definition of 'incipient talk'. Further work on the essence of 'incipient talk', including the comparison of dyads to multi-party interactions, could be of great importance in studying this phenomenon.

## **[6] Some Possibly Related Concepts**

Throughout modern interactional research, from traditions ranging from sociology to psychology to linguistics, authors have attempted to categorise types of interactions and define interactional frameworks. Some of the concepts that have been proposed touch on some of the same issues as Schegloff and Sacks's continuing states of incipient talk and subsequent uses by other authors. Two of these concepts originate with Erving Goffman and focus on the local interactional environment as a whole – *unfocused interactions* (Goffman, 1963) and *open states of talk* (Goffman, 1967). Another, *Gesprächsinseln* (Baldauf, 2002), focuses on the talk itself in relation to the ongoing interaction. We consider these concepts as possible alternatives in light of how the authors in our sample have used 'incipient talk'.

### *[6.1] Unfocused Interactions*

Goffman (1963, p. 24) defines unfocused interactions as 'the kind of communication that occurs when one gleans information about another person present by glancing at him, if

only momentarily, as he passes into and out of one's view'. It is worth noting that focused interactions do not have to solely involve talking to one another or have only short silences to progress unproblematically. He (ibid., p. 103) gives examples of Shetlanders who sit together and talk with pauses and gaps of several minutes whilst sitting around a fire, knitting, or smoking.

Goffman discusses focused interactions as an analytical default of sorts in which participants help to sustain a single focus of cognitive and visual attention (1967, p. 133)<sup>25</sup>. Whilst the visual attention may move from one participant to another as speakers change, interlocutors' focus is mutual at any given time. He contrasts such interactions with the aforementioned unfocused interactions, 'where individuals in one another's visual and aural range go on about their respective business unconnected by a shared focus of attention' (e.g., large social parties and street behaviour). Unfocused interactions involve people going about their business and engaging in talk as a matter of circumstance (e.g., in a queue). 'Focused interaction is the kind that goes on in a state of talk; unfocused interaction is the kind that goes on, say, when two persons size each other up while waiting for a bus, but have not extended to each other the status of co-participants in an open state of talk.' (ibid., p. 145)

Kendon (1988a) critiques the dichotomy of focused and unfocused interactions and gives several examples of each that meet some criteria of both. He suggests that rather than a dichotomy, there exists a continuum with focused and unfocused on opposite ends and the possibility for focused interactions to exist within larger unfocused interactions (as with individual parties within a queue, for example). He further distinguishes between a *common focus* (as in a queue) and a *joint focus* (as in a game) in addition to reiterating

---

<sup>25</sup> Stivers & Rossano (2010b) critique the treatment of talk-focused interactions as a point of reference, noting that most of day to day human interaction does not involve long stretches of talk.

Goffman's distinction between a *singly focused* (as in a business meeting) and *multi-focused* (as in a cocktail party) interaction.

### [6.2] *Open States of Talk*

In Goffman's *Interaction Ritual* (1967), he defines a 'state of talk' as one in which interlocutors 'have declared themselves officially open to one another for purposes of spoken communication and guarantee together to maintain a flow of words' (ibid., p. 34). Interlocutors may produce small talk in order to maintain a state of talk (ibid., p. 120). Although Goffman's use of 'open state of talk' in *Interaction Ritual* is ambiguous<sup>26</sup> and often simply refers to interlocutors being actively engaged in a state of talk, he later (1981, p. 104) describes 'open states of talk' as situations in which participants have 'the right but not the obligation to address remarks to the other participants, this being a condition that commonly prevails among individuals jointly engaged in a common task (or even similarly engaged in like ones) when this work situates them in immediate reach of one another.' However, the nature of initiating actions is such that participants are generally not obligated to produce them<sup>27</sup>, and again, as with Schegloff & Sacks, the description is conjectural rather than data-based.

A benefit of Goffman's 'open states of talk' is that it is sufficiently vague to not include, for example, all public transport situations. People are not usually confined to the proximity of an unwelcome speaker on forms of transport other than planes or cars. Nor does it assume that physical contingencies make interaction possible; it does not generalise across cultures. For example, it is not normative for English people to chat with fellow

---

<sup>26</sup> See Helm (1982) for a critique of Goffman's unformalised categories. Dynel (2011) formalises Goffman's usage of what he terms 'postulates on participant statuses in verbal interaction'. Although Dynel does not differentiate clearly between an 'open state of talk' and an 'unfocused interaction', these are, as we show, clearly different constructs despite Goffman's 'vernacular' usage of such categories.

<sup>27</sup> However, in ordinary conversation, any participant has the right to speak any time that a turn is not in progress and another participant has not been selected to speak (Sacks, et al., 1974). See also Schegloff (2007, chapter 10) regarding sequences of sequences, in particular action-type sequence series (ibid., pp. 207-213).

passengers on public transport, yet being seatmates is a widely cited situation that is described as a 'continuing state of incipient talk'. 'Open' is also useful, because it does not imply that talk necessarily happens in periodic chunks as *Gesprächsinselfn* (discussed in the next section) does. However, the phrase 'open states of talk' is talk-biased. By being talk-biased, it ignores the possibility of non-vocal practices being used to perform all or part of sequences (and therefore social interaction) (see also Berger & Rae, under review/Chapter 8).

### *[6.3] Gesprächsinselfn*

Baldauf (2002) uses 'Gesprächsinselfn' (which Gerhardt, 2009, translates as 'islands of talk') to describe groups of sequences of talk during watching football on television. In such a situation, interlocutors are at least acquainted with one another enough to know that they are interested in the events on screen, if not active enthusiasts. These are, again, similar to Goffman's open state of talk and could perhaps be said to constitute an anatomical part of an open state of talk (with another major component being the lapse in talk). Note, however, that 'open state of talk' does not imply anything about how talk is organised when it does occur.

During these *Gesprächsinselfn*, there is a greater orientation to spoken communication and less orientation to television-watching. This aspect of talk during television-watching would appear to be quite different in this respect from talk that occurs during physical activities such as surgery or other manual tasks in which the activity takes primacy even during speech production. However, like talk during such tasks, talk during television-watching may also be interrupted by events on television; participants may also alter their courses of action according to events on screen (Berger & Viney, under review/Chapter 6). One drawback of 'Gesprächsinselfn' is that it seems to ignore that monitoring of

interlocutors' vocal and non-vocal behaviour happens continuously rather than simply when talk occurs.

#### *[6.4] Confusion and Conflation*

Researchers have often used 'unfocused interaction', 'open state of talk', and 'continuing state of incipient talk' interchangeably or near interchangeably (e.g., Szymanski et al., 2006; Couper-Kuhlen, 2010; Pallotti, 2001). However, we see key differences in these concepts. Unfocused interaction refers to situations in which participants have not committed to a full-fledged interaction, whereas an open state of talk is one in which commitment has been made but in which talk is not required. Schegloff and Sacks's (1973) examples include situations in which interlocutors would be known to one another, 'members of a household in their living room, employees who share an office, passengers together in an automobile, etc.' This is more consistent with Goffman's open state of talk than with his unfocused interaction.

Another key difference between what Goffman describes as 'focused' and 'unfocused' and what Schegloff & Sacks describe as 'continuing states of incipient talk' is the process by which an interaction goes from one state to another. The transition from an interaction being unfocused to being focused involves a process in which, for example, interactants size each other up for conversation (Kendon, 1988a; Mondada, 2009). However, in a 'continuing state of incipient talk' such transition is taken to be instantaneous, with no processes to ease into talk with one another and the possibility for talk being omnipresent (Schegloff & Sacks, 1973). The latter would again appear consistent with Goffman's 'open state of talk' in that interlocutors have 'the right but not the obligation to address remarks to the other participants'.

Schegloff (2007, p. 26) describes 'continuing states of incipient talk' as situations in which 'participants are committed to co-presence by an event structure not shaped by the interaction itself ... may include familiars ... but can include strangers as well, whose juxtaposition is wholly incidental'. The latter would appear consistent with 'unfocused interactions'. Indeed, Couper-Kuhlen (2010) uses focused and nonfocused [sic] (citing Goffman) alongside 'continuing states of incipient talk' to describe similar data (see also Stivers & Rossano, 2010a). These uses underscore how ingrained and unquestioned 'continuing states of incipient talk' has become in CA. That one can conflate three concepts without backing such a combination up with data or even clear and deliberate conceptualisation is a failure in the field that needs to be remedied as we move forward.

'Continuing states of incipient talk' has been used as a buzz-phrase for challenging developments in conversation analysis, despite a lack of data-based analysis on the subject. Stivers & Rossano (2010a,b) begin to break down the assumptions behind our current conceptualisation of sequence organisation, suggesting that response relevance and engagement in spoken interaction lie along a continuum. Although Schegloff (2010) supports the idea of a continuum, he continues to dichotomise interactions. Both Couper-Kuhlen (2010) and Schegloff (2010) challenge Stivers and Rossano's (2010b) data-based analysis on the grounds that the contexts of the wide variety of interactions from which their data are drawn constitute 'continuing states of incipient talk' and assert that this accounts for their findings.

All of the terms proposed except for 'focused/unfocused interaction' are talk-biased and therefore ignore the possibility of actions such as writing answers or responding through non-vocal practices as constituting part of an ongoing interaction (see also Berger & Rae, under review/Chapter 8). They also ignore the possibility that silence is



communicative and can be used to prevent initiation of verbal communication, thereby maintaining social space (Saville-Troike, 1985). However, focused/unfocused interactions, as Kendon (1988a) has noted, are neither discrete entities nor well defined. The joint/common and single/multiple focus(-i) issues remain to complicate the model so that a continuum is actually on multiple axes and may become unwieldy for research that does not specifically look at the degree to which an interaction is focused or unfocused. Stivers and Rossano (2010a) similarly address the problem of dichotomising conversation (cf. Schegloff, 2010) along these lines. However, Philips (1985) describes interactions in terms of those that are structured through silence and those that are structured through talk. Although most interactions among humans involve a range of verbal, vocal, and non-vocal implementation, interactions or phases of interactions tend to be primarily structured around either silence or talk. That is, although, for example, assembling a desk together is overall structured through physical activity (or in Philips's terms, silence), talk may be occurring unrelated to the task or in reference to the task. Phases such as discussing the instructions may be structured through talk more than phases involving concentrated effort in aligning pieces, where talk merely accompanies a focus on physical activity. Philips notes that in interactions that are structured by talk, economy of speech and anaphora are much more common than in interactions structured through talk. This is particularly true with regard to the verb 'do', which can be used repeatedly without revival of the full verb phrase (Philips, 1985, p. 210). This avenue has been little-explored in CA and is worth further attention.

### **[7] Do 'Continuing States of Incipient Talk' Exist?**

The Oxford English Dictionary defines 'incipient' as 'beginning; commencing; coming into, or in an early stage of, existence; in an initial stage'. If taking a literal view of 'incipient talk' or more specifically 'continuing states of incipient talk', we would really be

dealing with the *onset* of talk (including greetings) rather than the *potential* for talk. This is the usage that nine papers in our dataset have actually used in a literal, non-technical manner. Given this misnomer and that the data that conversation analysts, interactional linguists, discourse analysts, and others collect does not usually include strangers sizing one another up for interaction, Goffman's 'open states of talk' may be a more appropriate term than either 'incipient talk' or 'unfocused interaction' for this data<sup>28</sup>.

Uses and definitions of 'incipient talk' and 'continuing states of incipient talk' vary greatly and are in some cases directly contradictory. These definitions as well as use without explicit definitions (which often simply cites Schegloff & Sacks, 1973, with the phrase 'continuing state of incipient talk') have been contradictory at best. For nearly 40 years, speculation in a concluding note of one classic paper has been treated as fact and canon, and a body of literature has been built up with little orientation to uniformity in, or data-based analysis of, the concept. Each author has used a somewhat different approach to 'incipient talk' based on the particulars of their data, and we are far from a universal understanding of 'incipient talk'. We propose that an effort to clarify the concept, or even to verify its existence, would need to involve large scale cooperation among researchers.

Silence, and in particular lapses, is a common aspect of interactions described as incipient talk. While we do not propose that silence or even lapses constitute incipient talk necessarily or even that incipient talk *must* contain silence or lapses (although it appears unlikely not to), studying interactions in which there are lapses, in particular long lapses, would seem an appropriate starting point in clarifying the issue of incipient talk. The addition of many varied data of interactions that are structured through silence but do not necessarily contain silence (in that the activities are accompanied by talk) would aid in

---

<sup>28</sup> Because there are no terms that are empirically grounded, we are not advocating simply a switch to another term that may be equally problematic despite being more clearly defined.

locating possible differences and regularities. All of this must of course be in reference to 'ordinary' conversation, or else an a priori assumption of these data as 'incipient talk' would remain.

We are not suggesting new terms or advocating use of other terms, as categorising interactions is not possible without first clarifying the issues involved through extensive research. It is necessary, first and foremost, to have as a basis for analysis a large corpus of proposed cases of states of incipient talk, or more broadly interactions with long lapses between sequences. Combining existing corpora from a variety of settings may be one way of initiating such a line of enquiry. Ideally, however, full time recording in multiple settings should be attempted to gain a more thorough understanding of issues such as topic carry over.

Efforts to use data in research toward clarifying the issue of incipient talk should begin with unmotivated looking in the first instance. A focus on the possibility of incipient talk would likely cloud the issues. Rather, aiming to describe the interactions should lead to greater depth of analysis. An additional aim might be to describe similarities and differences among interactions that have lapses. Some research questions that would be relevant to the beginning of a cohesive study of environments with (suspected) lapses include:

How do the other activities that people are doing affect the sequential environment?

How do people who are involved in different activities or settings initiate and close sequences, sequences of sequences, topics, spates of talk, etc.?

What happens during the lapses? why that now? is something else going on? is this really a lapse or an activity that is necessary to progress the interaction? When there is no accounting for a lapse, these might be the most interesting sites of analysis.

We hope that these questions may spark interest in further developments and collaborations that are necessary for moving forward. Unfortunately the present state of research involving potential continuing states of incipient talk does not support it as a cohesive construct. Alternative constructs do not capture all of the uses or definitions of continuing states of incipient talk and have their own limitations in terms of inclusion and exclusion. Further large-scale research to clarify how people perform and transition between interactional states is necessary to produce a unified theory of interactional states. In the following chapter, I will build on the present chapter by examining the sequential environments that develop in one context that is commonly thought to constitute a continuing state of incipient talk and show how a comprehensive conversation analytic approach to contextual features is more appropriate than application of theoretical constructs such as incipient talk.

## Chapter 6

### How Television Plays into Sequence Organisation: Orientation to Conversational Structures During Television-Watching<sup>29</sup>

---

<sup>29</sup> The version presented in this chapter was written by me. Rowena Viney has been involved in discussions, data sessions, and given feedback on the written work. The version that has been submitted for publication includes her as a co-author. It has also been presented as:  
Berger, I. & Viney, R. (2010). *Talking during films: Differing treatments of absent and gestural responses by first speaker*. 4th Conference of the International Society for Gesture Studies. Frankfurt (Oder), Germany.  
Viney, R. & Berger, I. (2010, July 8). *That's what I'm talking about: How television plays into sequence and topic organisation*. 3rd International Conference on Conversation Analysis. Mannheim, Germany.

**Abstract**

One environment that is commonly thought to constitute a continuing state of incipient talk is that of television-watching. This chapter seeks to address the extent to which structural organisations of conversation, namely sequence organisation and turn-taking, are affected by television-watching. Conversation analysis underpins the methodology, and the data consists of 16 hours of video recordings of familiars at home watching television or films, mainly in dyads. All participants are native English speakers and live in the United Kingdom or United States. We then discuss conversation analytic findings that are relevant to the current study. We show that television-watching affects certain aspects of sequence organisation but that other aspects of sequence organisation and turn-taking remain unaffected. We discuss television-specific orientation and offer possibilities for accounting for interactional differences between television-watching and other activities. A comprehensive conversation analytic approach to contextual features is more appropriate than application of theoretical constructs such as incipient talk. We conclude by offering a call for further research using larger data sets and hypothesis testing in accounting for the observed differences between interactions.

### **[1] Introduction**

One environment that is commonly thought to constitute a continuing state of incipient talk is that of television-watching. Empirical research on television-watching has often looked at how people relate to or are affected by the material on screen rather than how they interact with one another. The field of audience studies has formed around this research as well as other forms of media (see, for example, Brooker & Jermyn, 2003). However, audience studies is an inter-disciplinary field that has seen an increase in conversationally oriented research. A strong subfield exists around the activities of sports audiences and stylistic strategies of sports commentators (e.g., Kern, 2010). Likewise, work on situation comedies bridges research on humour in the media and interactional construction of targets of laughter (sometimes called ‘laughables’) such as jokes (e.g., Varney, 2012). This chapter seeks to address the extent to which structural organisations of conversation, namely sequence organisation and turn-taking, are affected by television-watching. Sequence organisation is the structure of conversation including initiating actions and responsive actions, whilst turn-taking is the process by which speaker transition and selection of the next speaker occurs.

The method that we use in this chapter is conversation analysis (CA), an observational technique with roots in sociology that involves sequential analysis of naturally occurring interaction that has been recorded either with video or with audio only. In this case, all of our data had been video recorded. Our dataset consists of 16 hours of video recordings of familiars at home watching television or films, mainly in dyads. All participants are native English speakers and live in the United Kingdom or United States. The data were collected as part of a larger project looking at interactions within households. The data were collected with the participants’ informed consent and have been pseudonymised for publication.

## **[2] Incipient Talk**

Couper-Kuhlen (2010) claims that during ‘incipient talk’, sequence organisation does not apply (See Stivers & Rossano, 2010a, for a response). ‘Incipient talk’ (or ‘continuing states of incipient talk’) refers to a hypothetical interactional state, which Schegloff and Sacks (1973) describe as one that does not require closing the interaction for resulting silences to not be attributable to a particular party. A complex and disjointed theoretical area has grown up around this phrase without empirical data as to the potential nature or homogeneity of such interactions (Berger, Viney, & Rae, under review/Chapter 5).

Interactions that have been described as ‘incipient talk’ or ‘continuing states of incipient talk’ have included environments as varied as push-to-talk radio use (e.g., Szymanski, et al., 2006), airline ground control (Suchman, 1997), and operating theatres (Mondada, 2006). Some commentators at conferences where we have presented have suggested that our data constitute ‘continuing states of incipient talk’ and may have a turn-taking (cf. Sacks et al., 1974) and/or sequence organisational (cf. Schegloff, 2007) structure of their own.

This assertion, however, has multiple difficulties. 1) Uses of the phrases ‘continuing states of incipient talk’ and ‘incipient talk’ have varied dramatically across authors (Berger et al., under review/Chapter 5), 2) No empirical research has shown that ‘continuing states of incipient talk’ exist as a participant-oriented to state (Berger et al., under review/Chapter 5), and 3) Participants in our data indeed orient to both turn-taking and sequence organisation as in other environments that have been studied. Certain types of initiating actions and turn designs have less response relevance than they do when parties are engaged primarily in talk. That is, although they make relevant responsive actions, they do not require them. This holds true even when turn designs have conditional relevance (a response is required) in ordinary conversation. This, and not the overarching sequence organisational structures, is what can be said to be context-specific.



Berger et al. (under review/Chapter 5) have shown that existing concepts such as incipient talk (Schegloff & Sacks, 1973) that purport to account for differences from canonical conversation are inadequate when confronted with actual naturally occurring interactional data. In a systematic review and content analysis, they show that such concepts are furthermore ill-defined and have never been held up to empirical scrutiny, and that three divergent definitions and usages of ‘incipient talk’ predominate. They suggest that only large-scale analysis of individual contexts of proposed ‘incipient talk’ can begin to clarify any possible definition, boundaries, or properties. ‘Incipient talk’ originated as a brief speculative comment by Schegloff & Sacks (1973) in their foundational work on closings in conversation. It has since been used without empirical grounding, and uses and definitions have varied widely.

Some conversation analysts have begun using the terms ‘focused’ and ‘unfocused’ to describe interactions (e.g., Couper-Kuhlen, 2010; Mondada, 2009; Szymanski et al., 2006; vom Lehn, Heath, & Hindmarsh, 2001). Goffman (1963, p. 24) describes ‘unfocused’ interactions as ‘the kind of communication that occurs when one gleans information about another person present by glancing at him, if only momentarily, as he passes into and out of one’s view’, which Goffman originally used to discuss the kinds of interactions that people have in public with strangers, for example at bus stops. It has since been extended and used to discuss a number of different situations involving strangers and non-strangers in co-present interaction (Kendon, 1988a). However, there are similar difficulties and context-specific issues in the use of the terms. For example, is a medical examination to be considered focused or unfocused? It cannot be compared simplistically to either a tea party (in which participants are primarily engaged with talk) or an oil change (in which participants are primarily engaged in manual tasks). Even in the case of the former, talk is

not the only event occurring, and interactants may oscillate between focus on talk and focus on drinking, eating, cleaning spills, and getting more food or drink. Interactions often do not have one purpose, either. Despite being ostensibly task-oriented, helping a friend assemble furniture may be done not only to assemble the furniture but also to catch up on each other's lives or otherwise spend time together. Moreover, interactions may have multiple foci on different levels (Kendon, 1988a). For example, a queue – which has all members focused on, for example, buying tickets, but which contains many parties engaged in their own conversations and not with others' – does not fit neatly into the focused/unfocused dichotomy.

Schegloff & Sacks (1973) put 'incipient talk' onto the table wholly theoretically in contrast to talk that ends with closings and needs to be restarted with greetings when talk resumes. In other words, 'continuing states of incipient talk' referred to situations that had some degree of lapse between sequences rather than separate interactions. However a) continuous talk takes quite a bit of work to sustain b) others treat 'incipient talk' with a variety of conceptualisations including responses not being conditionally relevant (Couper-Kuhlen, 2010). We argue that a) tolerance for lapses is context specific b) orientation to sequence organisation remains when not continuously talking (i.e. conditional relevance remains) c) topics and frequency of talk may depend on factors such as how much attention tasks require, whether tasks require talk to coordinate, whether tasks are cooperative or individual (including whether tasks are done by both parties). At this time detailed analysis in a variety of contexts is needed to begin to build an empirically grounded theory of states of talk. We do not claim to provide such a theory but rather to contribute a detailed analysis of television-watching to this endeavour. Because there is no clear definition of incipient talk, there is no way to place television-watching definitively

in or out of this category. Rather than attempt to create yet another definition, we approach the current data from how participants themselves behave.

### **[3] Sequence Organisation & Turn-taking**

We thus aim to provide an account of how sequence organisation (and by extension, turn-taking) is affected (and also *not* affected) during television-watching using conversation analytic methodology. Here we will review basic findings in CA that are relevant to the phenomena under examination.

#### *[3.1] Adjacency Pairs*

Sequences of talk are relatively discrete units enacted by two or more participants. The basic form is that of an adjacency pair (Schegloff, 1968) that consists of two turns, a first pair part (or initiating action) and a second pair part (or responsive action). Adjacency pairs may be expanded upon to result in lengthy sequences. However, they remain the basic sequential unit even in lengthy sequences (see Schegloff, 2007 for a thorough primer on sequence organisation). First pair parts are regularly responded to with second pair parts, although some require a response (make a response conditionally relevant) and others simply make a response relevant but not required (Schegloff, 2007, p. 20). Such responses should be type-conforming and type-related – they should be in the implicated format and do the implicated action. For example, a yes/no interrogative regularly receives a yes/no response. Stivers & Rossano (2010b) have also shown how first pair parts that do not per se make responses conditionally relevant (in their case, assessments – explicit evaluations of situations, ideas, people, or objects) can be designed to increase the likelihood of a response (cf. Couper-Kuhlen, 2010; Schegloff, 2010; Stivers & Rossano, 2010a).

#### *[3.2] Speaker Selection and Turn Allocation*

In interactions with only two people (most of our data in fact consists of dyads), the expected next speaker after a completed turn is usually the participant who is not currently speaking. In groups of three or more, features of the turn such as relative epistemic authority of recipients, use of names or other address terms, and speaker eye gaze or gesture can display who is selected to speak next. (Sacks et al., 1974) Silences that occur between turns within a sequence (gaps) can be attributed to a particular participant, i.e. that a turn is relevantly absent and that the absent turn belongs to a particular person (Sacks et al. 1974; Schegloff, 2007). Such a gap after a first pair part can signal that a dispreferred second pair part is forthcoming (Pomerantz, 1984a; Sacks, 1987; see also Schegloff, 2007, on preference more generally). ‘Dispreferred’ refers to actions that are generally avoided in conversation, whether by delaying or mitigating them (and without elaboration often come across as curt or impolite). Dispreferred responsive actions include, but are not limited to, refusing invitations or offers, answering negatively to questions, and disagreeing with assessments. In such a situation, the speaker of the first pair part may produce additional utterances to elicit a preferred response.

#### **[4] Orientation to Existing Structures**

As we have mentioned previously, some have proposed that interactions described as ‘incipient talk’ lack conditional relevance in general (Couper-Kuhlen, 2010). However, Stivers & Rossano (2010a,b) have shown that for assessments at least, the requirement to respond can depend on the how the first pair part is designed. In our data (which involves a situation often regarded as ‘incipient talk’), participants orient to both turn-taking and sequence organisation in ways that are similar to interactions in other (non-incipient talk) environments that have been studied. Despite a similar overall orientation to social interactional structures, certain types of initiating actions and turn designs have less response relevance during television-watching than they do when parties are engaged

primarily in talk. This relaxed response relevance, and not the overarching sequential structures, is what can be said to be context-specific (and indeed this appears to vary with the level of focus of participants on the television vs. each other).

#### *[4.1] Adjacency Pairs*

Initiating actions in our television-watching data regularly fail to elicit responses typical of ordinary conversation. When an initiating action is produced, it makes relevant (but does not necessarily require) a response. Certain turn designs may be more likely to receive a response. For example, the majority of yes/no interrogatives in our data are responded to by participants. However, the response may be quite minimal compared to that which would be expected during ordinary conversation (see Extract 1, below). Responses tend to include a) no response, b) a vocalisation in response (often quiet and short), or c) a non-vocal response (not always obviously visible to the first speaker<sup>30</sup>). This is in contrast to ordinary conversation in which talk responses are overwhelmingly common (and if not talk, at least a gesture or vocalisation).

#### *[4.2] Minimal Responses*

Gestural responses are more common in television-watching than in other interactions with the same participants. Gestures are situated in sequentially relevant places and are generally minimal (e.g., a single nod) and type-conforming (see Berger & Rae, 2012/Chapter 7, regarding uses of non-type conforming gestures). Extract 1 represents one such a gesture.

Extract 1 [CA 052309]

---

<sup>30</sup> Although the interlocutor might not be gazing at the one performing the action, it is sometimes possible that the non-vocal practice could have been seen in their peripheral vision. However, movement in the peripheral visual field tends to evoke the fixation reflex - eye movements that put the source of movement in the central visual field (Arden, 1997, p.34). None of the participants in our data are known to have peripheral vision disorders that would disrupt this reflex.

01 Cor: Did she say Charlie ((looking at Al))  
02 Al: >> ((nod))  
03 Cor: I thought that was her dad

Here Coral asks Al a question about an event that just occurred on screen (line 1) and is gazing at him as she produces her turn. Al responds with a single vertical head nod (line 2). Al's nod signifies an affirmative response, and Coral then produces a turn 'I thought that was her dad' that is responsive to Al's confirmation (line 3).

Similarly, minimal vocal responses occur in sequentially relevant places and use prosody to convey their meanings (Gardner, 1997; cf. Wiggins, 2002). An advanced, digital prosodic analysis is beyond the scope of this chapter, and we offer the latter observation as culturally competent hearers and based on how participants treat the utterances. Extract 2 is an example of a minimal vocal response.

Extract 2 [Viney, DN4]  
01 Daw: Man she was really worried about him.  
02 (0.4)  
03 Nil:>> Hn

Dawn issues an observation in line 1 about events on screen, and after a 0.4 sec. delay, Niles responds with 'hn'. His response closes the sequence, and there is a lapse before the next sequence.

#### *[4.3] Preference*

Some actions more strongly implicate a response and, more specifically, a response of a certain type. For example, yes/no interrogatives strongly implicate answers in the form of yes or no (Raymond, 2003). For most yes/no interrogatives, yes is the preferred response

(i.e., the response comes quickly and is not mitigated). On the other hand, assessments do not strongly implicate a second assessment, but second assessments regularly occur. A preferred response for an assessment is not simply agreement but an upgraded second assessment (Pomerantz, 1984a). Any other response, even equal agreement, is treated as disagreement.

In our television-watching data, second assessments are uncommon, and responses to assessments usually come in the form of acknowledgement tokens using prosody to convey agreement (as in Extract 2, above). Canonically upgraded second assessments that include stronger wording are nearly absent, with the most elaborate second assessments taking forms similar to Extract 3.

Extract 3 [Viney, DN4]

01 Nil: °>Ah e look< so young, °  
02 Daw:>> Yeah he does doesn't he.

In Extract 3, Niles produces an assessment of a character on screen 'ah e look so young' (line 1), to which Dawn replies 'yeah he does doesn't he', rather than for example 'he looks like just a kid' as might be expected of an *agreeing second assessment* in ordinary conversation. Dawn's formulation of her agreement with a tag question 'doesn't he' (line 2) marks her assessment as thought through in response to Niles's and demonstrates that she has come to the same conclusion. Although she does not upgrade her assessment in the classical sense, she makes her agreement clear.

Whilst not always assessments in the technical sense, joking about and mocking characters on screen may receive *relatively* enthusiastic responses regardless of whether they are produced in the form of assessments or imitations. Extract 4 involves Al mocking

a character who has asked a teenage woman whether she remembers him playing Santa Claus when she was a young child (lines 1 and 7).

Extract 4 [CA 052309]

01 Al:     ↓He don't know her:::  
02           (1.5)  
03 Cor:    Oh the the Santa Claus  
04 Al:     Yeah  
05 Cor:    huhuhuh  
06           (1.9)  
07 Al:    >> Hey Bella remember me (.) four years old come on now  
08           (.)  
09 Cor:>> Duh

After repairing the trouble source of 'he' in line 1 (lines 3-4), Coral laughs (line 5). Al then produces an imitation (line 7) to which Coral responds 'duh' in line 9. Coral's 'duh' further mocks the character and displays agreement with line 7.

#### **[5] A Television-Specific Orientation: Television as a Topic Resource**

Topics in our data tend to be stimulated by on-screen events. When non-television/film-related topics are introduced, they tend to follow on from related television-stimulated talk. For example, Extract 5 involves a participant initiating the topic of mutual friends through a character on the television who reminds her of an acquaintance.

Extract 5 [Viney, KB 'Steve']

01 Bar:    >°Mp I really don't like him°.<  
02           (.)  
03 Bar:    But he reminds me of (0.2) Steve. you know the one  
04           who came to new ye:ars.



05 (1.6)  
 06 Kim: No(h):, hh [heh heh]  
 07 Bar: [Lisa's ] fiancé.=  
 08 Kim: =Oh yeah:.  
 09 (0.4)  
 10 Bar: (N- n- [e's real-] )  
 11 Kim: [Why does he r]emind you of °Ste[:ve°. ]  
 12 Bar: [E's got]  
 13 the same kind of (0.4) charismatic: (0.4) (ness) (.)  
 14 (in) the way >that e< talks an s[tuff.]

In line 1, Barbara refers to a character on screen 'I don't really like him' and continues her turn in line 3 'but he reminds me of Steve ...'. Barbara and Kim establish the identity of the acquaintance being referenced as 'the one who came to new years' (lines 3-4) and 'Lisa's fiancé' (line 7). Kim asks in line 11 why the character reminds Barbara of Steve, and Steve's personality is topicalised (lines 12-14, continuing data not shown). They then further discuss mutual friends leading off from this topic (data not shown).

Extract 6 shows a similar treatment of the television as a resource for topic initiation. However, in this extract, the adjacency pair is the only immediate talk on the topic. In Extract 6, Dawn and Niles have been watching the film *Y Tu Mamá También* with English subtitles. Music begins playing on screen, and Dawn asks Niles whether he has the soundtrack for the film (lines 1-2). He shakes his head (line 4) before very quietly saying 'no no' (line 5). There is no further discussion for the immediate future, as shown in line 6 by a 7 second lapse. The topic is later taken up again with the participants discussing the bands on the soundtrack (data not shown).

Extract 6 [Viney, DN4]

01 Daw:>> ((points to TV while looking at Niles)) D'you have  
 02 >>the sound °track for this°.

03 (0.2)  
04 Nil: ((shakes head three times over 0.2 seconds))  
05 Nil: °°No. no.°°  
06 (7.0)

Participants may treat the television as a 'third entity' that is a resource for topic but that can also be addressed as a non-participating participant. In Extract 7, Coral, who is oriented to the screen, speaks directly to the character about what is happening (lines 1 & 3-4).

Extract 7 [CA 052309]

01 Cor:>> Yes Bella he was there  
02 (1.1)  
03 Cor:>> But he disappeared before you could get the light  
04 ohn

Obviously, Coral is not expecting a response from Bella, however, that is not to say that speaking to the screen never has interactional implications. Examination of the sequential environment reveals an immediately prior case of pursuit that has been left un-responded to (Extract 8).

Extract 8 [CA 052309]

01 Cor: Uh oh it's a full moon  
02 (2.7)  
03 Cor:>> You know what happens when a full moon comes out  
04 don't you ((turns to Al and back during the turn))  
05 (12.0)  
06 Cor: Yes Bella he was there  
07 (1.1)  
08 Cor: But he disappeared before you could get the light

09 ohn

10 Cor: ((smiles and turns to Al then back to the TV))

In line 1, Coral produces a comment on the state of events on screen, projecting upcoming negative events ('uh oh it's a full moon'). Al does not respond, and a 2.7 second gap forms (line 2). Coral (lines 3-4) does pursuit by addressing Al directly, 'you', expanding '... what happens when a full moon comes out ...', and adding a tag question 'don't you' and gazing at him. Al still does not respond, and Coral attempts no further pursuits, keeping her gaze at the screen until line 9. However, in lines 6-9, she addresses the character on screen 'yes Bella he was there' and 'but he disappeared before you could get the light ohn' before smiling, turning to Al, and then turning back to the television (line 10). Coral's sequence with Bella (lines 6-9) appears to serve a rather cheeky function of continuing to talk despite it being unwelcome (Al has also previously rebuked her for talking, See Section 7.3). Al performs watching the television, rather than engaging in conversation, through his silence. Although Al is the only possible real-time and responsive recipient of Coral's talk, by ostensibly talking 'to the screen', she has negotiated a clever way to continue talking for his benefit. She is in effect not talking *to* Al yet still talking within his range of hearing.

## **[6] Actions**

### *[6.1] Triggers for Initiating Actions*

Initiating actions are actions that begin sequences, for example offers that make relevant responsive actions such as acceptance or rejection. They may be part of a larger spate of talk or 'out of the blue'. The initiating actions in our data tend to be related to topic initiators that are stimulated by on-screen events and the subsequent initiating actions within those topics. This means that the majority of initiating actions in our data are in some way connected to on-screen events, even if not directly. A notable exception is that

of environmental factors such as expressing physical discomfort or requesting an item from another participant. For example, in Extract 9, Coral topicalises the species of a bug that she sees across the room.

Extract 9 [CA 062010]

01 Cor:>> (Looks like     ) a giant mosquito flyin around today  
01            (5.4)  
02 Cor:    Can you see if that is a moth or  
03 Al:      What  
04 Cor:    The bug flyin around the kitchen is that a moth  
05 Al:      Yeah  
06 Cor:    Oh okay

Unlike other extracts we have examined, Extract 9 involves three participants. In line 1, Coral makes a statement about the bug, and this receives a response from neither Al (who is seated next to Coral) nor Ernie (who is across the room near the bug) (line 2). Coral pursues a response at line 3 with a question designed to elicit the identity of the bug. It is apparently directed to Al through Coral's non-vocal behaviour but could be easily heard and answered by Ernie as well. Coral's question is probably headed to 'moth or mosquito', but in line 4, Al produces an open class repair initiator (Drew, 1997) 'what' just after Coral's 'or' at line 3. Coral issues the repair in line 5 as a yes/no interrogative with the full subject in first position, 'the bug flyin around the kitchen is that a moth'. Al responds affirmatively in line 6, and in line 7, Coral ends the sequence receiving the information with 'oh okay'. Initiating actions with no clear trigger are rare in our data of television-watching as the primary activity (as opposed to eating dinner with the television in the background).

Although in ordinary conversation, initiating actions routinely receive responses, initiating actions in our television-watching data regularly are not responded to or are minimally responded to. A characteristic sequence (or rather initiating action alone, as the case may be) is shown in Extract 10, in which Niles produces an assessment ‘(this/that) place looks so cool’ and receives no response from Dawn, who maintains her bodily orientation toward the television.

Extract 10 [Viney, DN4]

01 Nil: (This/that) place looks so cool. huh huh

02 (18.2)

Bodily orientation and *doing* being busy watching/listening may contribute to the acceptability of non-response and minimal response. Whether the speaker of the initiating action does work designed to get a response (e.g., address term, turning to recipient, tag question) can make a response more likely (see also Stivers & Rossano, 2010b).

Berger & Rae (2010) describe how actions can be built non-vocally and can be used to do actions that talk cannot in a given situation. Non-vocally-implemented actions are much more common in television/film watching, and by engaging in gestures during television/film watching, interactional trouble related to problems with hearing or speaking are largely avoided. We propose that they can be used both to avoid talk overlapping with what is on screen in any given instance but also to limit engagement in conversation during the show or film. Gestures may even be used when participants are unlikely to see the gesture. Whilst we cannot determine definitively whether participants are able to see each others’ gestures, Extract 11 provides a particularly clear example of a case when gestures are used and it is highly unlikely that the other participant was able to see them. Extract 11 involves finger-waving being used to replace a word in an *initiating* action. Again, this

extract is being used to demonstrate the use of gesture *at all* when it is unlikely to be seen, rather than specifically responsively.

Extract 11 [Viney, DN4]

01 Daw: It's weird how it- [(d-) ] yeah.  
02 >> [(waves fingers)]  
03 >> (0.2) ((Niles shifts gaze to Dawn))  
04 Nil: [ɛy(h)eahf. ]  
05 [(Dawn slight head nod)]  
06 Nil:>> ((Niles shifts gaze to screen))  
07 Daw:>> ((waves fingers))  
08 (33.0)

The sound in this film occasionally cuts out abruptly in order for the narrator to speak over the scene. There tends to be silence between the sound of the scene stopping and the narrator's voice beginning, and it can have the appearance of a problem with the audio. The scene that Dawn and Niles are watching has just cut out in order for the narrator to speak, and Dawn comments in line 1 about it being weird. She begins 'd-' (possibly going toward 'does that') in line 1 but instead waves her fingers (line 2). Her talk-gesture hybrid turn constitutes a complete action (Olsher, 2004). Niles shifts his gaze (line 3) to Dawn and says 'yeah' (line 4). At the same time, Dawn gives a slight nod (line 5). Niles's 'yeah' at line 4 demonstrates that he understands and agrees. Dawn, having never produced the lexical item she was searching for, nods. This nod serves to confirm that she understands that Niles understands what she meant. Niles then shifts his gaze from Dawn to the screen (line 6), and Dawn gives another finger-wave gesture (line 7). Although it is possible that Dawn might have thought that her gesture could be seen at lines 1-2, Niles has had to move his head to gaze at her (line 3) and subsequently moves his head again away from her (line 6) between the first and second gestures. Nevertheless, she produces

the same gesture in line 7. This is perhaps further closing the word search that she had begun at line 1, as the finger-wave has rather dismissive connotations in British and American English.

Gestural and absent responses are both treated as unproblematic. However, differences exist between them, showing that gestural responses are indeed treated as responses. Gestural responses may receive a response from the first speaker that recognises the sequential relevance of the gesture (as in Extract 1). Absent responses, however, may result in the first speaker orienting back to the screen or (in some cases) topic pursuit.

### *[6.3] Pursuit*

When a response does occur, gaps tend to be no lengthier than they are during other interactions involving the same participants (but see Berger, 2011/Chapter 3, on contextual considerations in analysing silence). Two major exceptions to this are when there is action or unusually loud volume on the television. When a response is delayed, participants do not often orient to a possible dispreferred response by re-framing the initiating action. Similarly, minimal responses and non-responses do not tend to be treated as problematic or inadequate by the speaker of the initiating action. Pursuit refers to actions that are taken by the speaker of an initiating action when a response is not forthcoming or a dispreferred response appears to be in the works. On occasion, pursuit occurs following non-response, but it is abandoned if no response is forthcoming. We have no cases of pursuit following gestural or minimal responses, indicating that they are treated as acceptable responses.

Lack of response can indicate trouble in the conversation, and pursuit is designed to re-establish alignment between parties. However, in our data, lack of a response even to pursuit tends to come off unproblematically, with the speaker simply dropping the

sequence rather than escalating the pursuit or rebuking the non-responder. Extract 12 (an abridged version of Extract 8) is one such an example.

Extract 12 [CA 052309]

01 Cor: Uh oh it's a full moon  
02 (2.7)  
03 Cor:>> You know what happens when a full moon comes out  
04 don't you ((turns to Al and back during the turn))  
05 (12.0)

As noted in our discussion of Extract 8, this extract leads to another sequence which may be related to Al's previous rebuke of Coral and his associated disengagement in talk during the film. However, the later talk by Coral to the character on screen does not come for a full 12 seconds after Coral's pursuit and turning to the television at lines 3-4. Similarly, Extract 13 (an extension of Extract 1) shows a turn that receives no response (line 1), a 3.2 second gap (line 2), a pursuit (line 3), continued silence (line 4), and turning back to the television (line 5).

Extract 13 [CA 052309]

01 Cor: I thought that was her dad  
02 (3.2)  
03 Cor:>> Or she's just callin her dad Charlie  
04 (1.9)  
05 Cor: ((turns to TV))

This pattern regularly occurs with both pursuit and initiating actions in which participants turn to face the recipient. Pursuit (if it occurs at all) is abandoned following non-response to it, and the speaker orients back to the television. Whilst participants may relax their expectations of a response (evinced by lack of either a response or pursuit



following a first pair part and lack of sanctions on participants for not responding), they still orient to the overarching structures that make a response a relevant next action and when that action should occur, if the responsive action is produced at all.

## [7] Accounting for the Differences

### [7.1] Interactional Focus

Although ‘incipient talk’, ‘focused’ & ‘unfocused’ interactions, and other terms do not have well defined boundaries, there appears to be a multi-dimensional spectrum at least for focused/unfocused interactions (Kendon, 1988a). Table 1 illustrates extremes of three kinds of focus (unfocused interactions are undefined, as all interactions have some degree of focus).

Table 1. Kendon’s (1988a) dimensions of interactional focus			
Dimension	common focused	jointly focused	multi-focused
Examples	television, parade	conversation, game	cocktail party

Kendon also points out that interactions may be of different degrees or types within others. For example, a queue may consist of multiple jointly focused interactions (conversations) within a larger common focused interaction. A set of queues could even be a multi-focused interaction that contains common focused interactions that contain jointly focused interactions. Additionally, foci may be of relative ‘strength’. In the data we have presented thus far, television is the main focus of the interaction. Compare this to talk that occurs in interactions where television is merely *a* focus, such as watching television in the background during dinner. Whilst our dataset is too small to provide useful statistics about the relative amount of vocal interaction across types of interaction, some indication of differences may be gained from comparing the transcripts of interactions ‘typical’ of our

dataset. A typical transcript for 40 minutes of television-watching may be 2500 words, whereas a typical transcript of eating dinner whilst watching television in the background may be 4500 words. Approximately double the talk occurs during interactions in which television is merely a focus rather than the focus.

In interactions in which television is merely a focus or not present at all, actions are regularly responded to and much more elaborately so than those in which television is the main focus. CA has explored the ways in which talk is structured in ordinary and institutional conversation in great detail (see e.g., Schegloff, 2007). Interactions in which television is not the primary focus appear much closer to those canonically described by conversation analysts. For example, Extract 14 is taken from a dinner between Coral and Al in which they have the television playing in the background. Coral is clearing the table, and Al is seated at the table. He is turned toward the television. Immediately prior to this extract, they have been discussing their plans for the evening.

Extract 14 [CA 052109]

01 Cor: Do you have an early day tomorra or a late day  
02 (4.4)  
03 Al: Ah go in at eight o'clock  
04 Cor: I call that an early day then  
05 (1.2)  
06 Al: Nah I call that a regular day (.) early day is  
07 when I go in at six o'clock[ an] I have to get up=  
08 Cor: [O:h]  
09 Al: =and do [that  
10 Cor: [When that's true I guess you can look at it  
11 that way >th' reason I said early day was because  
12 you usually go in early gettin off early instead a closin

In line 1, Coral asks Al whether he has an early or late day at work tomorrow. After a 4.4 second gap (line 2), Al responds that he has to be at work at 8 in the morning (line 3). Coral gives her interpretation of 'eight o'clock' as an 'early day' at line 4, and after a 1.2 second gap at line 5, Al responds that it is in fact a regular day to him (line 6) and that an early day is when he needs to be at work at 6AM (line 7). Coral receipts this with 'Oh' at line 8 in overlap with Al's continued talk 'an I have to get up and do that' (lines 7 and 9). Coral agrees with Al and accounts for her reasoning of 8AM being an early day (lines 10-13), beginning in overlap (line 10) with Al's 'that' at line 9. Such an exchange is quite common, and despite Al watching the television, the conversation is much more *conversational* than when the participants are engaged in watching television as an event. Although there are two fairly long gaps (lines 2 and 5), there are also two instances of overlap (lines 7-8 and 9-10) as well as extensive post-expansion (lines 4-12) of the base sequence (lines 1-3). Although multiple foci are present, it is not simply the presence of the television that affects how they engage with one another. It is through their talk that Coral and Al here perform having a conversation, rather than watching television. Additionally, watching television in this instance is not 'an event'. It is something that is happening in the background to their dinner, and they orient periodically to watching it, talking about it, and talking about other matters.

When there is a delay in responding, speakers of first pair parts often pursue a response in these interactions in which television is one of multiple foci or not present at all. Pursuit tends to be much more successful in achieving a response than when television is the main focus. Extract 15 is taken from the discussion immediately preceeding Extract 14, in which Coral and Al are discussing their plans for the evening. Coral is behind Al at the kitchen sink, whilst Al is at the dinner table turned toward the television.

Extract 15 [CA 052109]

01 Cor: So what's my love doin t- the rest of the night  
02 (6.0)  
03 Cor:>> H:uh?  
04 (3.2) ((Al turns slightly toward Coral))  
05 Al: Workin on some craft stuff and (1.2) showin you  
06 stuff on the computer

In line 1, Coral asks Al what his plans are for the rest of the night. After a 6 second gap (line 2), Coral asks simply 'huh' (line 3), and after a 3.2 second gap (line 4), Al produces an answer (lines 5-6). His answer is quite elaborate compared to many of the responsive actions in our television-focused data, as is characteristic of our non-television focused data.

Gestural responses are also much less common during interactions in which television is just one of multiple foci or not present at all. In these interactions, speech does not compete with the television in the same way that it does when the television is being actively watched. Gestures during interactions in which television is not the main focus also tend to be much more elaborate, perhaps because there is less visual competition. There is very little gesticulation (i.e. co-speech hand gestures) during television-watching compared to instances in which the television is not the main focus. Likewise, for responsive gestures, compare Al's single nod in Extract 1, during which the participants are engaged in watching a film, with his bowing gesture in Extract 16, both of which are positively valenced and in response to yes/no interrogatives. Coral and Al are facing one another at the dinner table with the news on television in the background.

Extract 16 [CA 052109]

01 Cor: Did you like your dinner Al

02 (1.4)  
03 Al: >> ((puts hands together))  
04 Cor: [ mean it looks like it  
05 Al: >> [ ((bows))  
06 Al: >> ((bows again))  
07 Cor: Well I'm glad you like it

Berger and Rae (2012/Chapter 7; under review/Chapter 8) have examined gestural responses to talk, including Extract 16, in detail. Here we wish only to demonstrate the relative complexity that both gestures and utterances have in non-television-focused interactions compared to during television-watching. A similarly talk-focused segment of an interaction during an interaction that is television-focused on the whole also involves the use of elaborate gestures to convey information. In Extract 17, Kim and Barbara are seated on the sofa but have been engaged in extended talk on a topic that we have previously discussed in Extract 5. Kim uses a gesture mimicking Steve's use of a mobile telephone in constructing her question clarifying to whom Barbara is referring (lines 5-8). Barbara then does a similar but more elaborate form of Kim's gesture that involves mimicking Steve's posture and holding the 'phone' to her ear (lines 12-13). Kim's initial gesture comes approximately 25 seconds into the topic of Steve, and Barbara's occurs at approximately 30 seconds.

Extract 17 [Viney, KB 'Steve']

01 Bar: E's got the same kind of (0.4) charismatic: (0.4)  
02 (ness) (.) (in) the way that e talks an s[tuff.]  
03 Kim: [Is h]e  
04 th(h)e (h)one th(h)at p(h)ut h(h)is l(h)eg up on the  
05 ta(h)bl(h)e .hhh whi(h)le he \*was talking on his  
06 mobile when you met Sarah  
07 [for the first time. (0.6) \*]

08        >> ((\*K mimics Steve using phone\*))  
 09        [((B is smiling and nodding))]  
 10 Bar:    Hn hn hn hn Yeah.  
 11 Kim:    That's dis[gusting.]  
 12 Bar:>>        [(        )] ((B mimics Steve holding the phone))  
 13        (1.0)  
 14 Kim:    [HEY::        ]  
 15 Bar:    [And that time when] like I was on the munch with  
 16        um ((sniff)) mph Sarah and Josslyn (0.4) and somebody  
 17        called me and I was like .hh excuse me ladie:s,

Kim's gaze during her own gesture is toward Barbara. However, following her turn, her gaze returns to the television. Barbara's gesture, although mocking Steve and fairly elaborate, is very short-lived. The time that elapses from the onset of her transition to Steve's posture to the completion of her transition to her original posture is approximately one second. Her imitation of Steve occurs in first position in a new sequence of actions related to the ongoing topic but constitutes what de Fornel (1992) has termed a *return gesture*. It picks up on major themes or forms of a prior gesture to construct a related action. Return gestures serve to increase intersubjectivity and demonstrate alignment with the prior gesturer (Koschmann & LeBaron, 2002). In this case, it also serves to demonstrate affiliation with Kim's negative evaluation of Steve. Because Kim was physically oriented to the screen when Barbara imitated Steve, it is possible that she did not fully see Barbara's gesture. Following one second of silence, both Barbara and Kim begin speaking again, and the topic of Steve continues.

### [7.2] *Watching as an Event*

The 'event' of watching a film or specific show may take priority over interaction. Gerhardt (2006) has looked at how viewers of televised football matches, in which

‘watching the game’ is the purpose of their gathering, organise talk around the actions on screen. She notes that talk tends to occur in short bursts when on screen action is low and refers to these as ‘islands of talk’ (or ‘Gesprächsinselfn’, see also Baldauf, 2002). We have demonstrated how participants can disengage from talk during television-watching and their resources for doing so. The preparation often involved in watching a film in particular (e.g., gathering or cooking snacks, putting a DVD in the player) and the extended attention needed to understand the overall story, may contribute to the ‘event’ of watching.

### *[7.3] Sensory Limitations*

Another reason for decreased vocal and non-vocal interaction whilst watching television or films may be due to human limitations in processing stimuli. As we have noted, gaps tend to occur when a first pair part has been produced shortly before or during loud or action-packed events on screen. Cognitive research has looked at relative attentional capacities when processing concurrent stimuli and performing concurrent tasks and supports this proposal. Talsma, Doty, Strowd, & Woldorff (2006) found that attentional capacity for processing concurrent stimuli was greater across modalities than within modalities. This accounts both for decreased interaction generally as well as increased gestural responses during television-watching. Concurrent processing of auditory stimuli (participants’ talk and on screen talk) is difficult, resulting in less talk. Because first pair parts tend to be spoken (cf. Berger & Rae, under review/Chapter 8), the range of relevant responses is limited and therefore provides a framework for understanding responses that do occur. Responses tend to also economise on auditory stimulation, taking the form of minimal vocalisations and gestures. This further reduces competing stimuli and contributes to participant focus on the events on screen. By engaging in gestures, participants make use of multiple modalities to decrease cognitive load and human capacity to maintain attention across multiple modalities.

*[7.4] Acceptability and Sanctioning*

Whilst we have no cases of a first speaker sanctioning a non-responding recipient, we have multiple cases of recipients sanctioning speakers of first pair parts. A first speaker may be sanctioned for ‘too many’ initiating actions or pursuits, although this does not necessarily need to be done in those terms. For example, Extract 18 involves Al sanctioning Coral after several un-responded to first pair parts, targeting her commentary on the film as problematic (rather than as too many comments or too much talk). Just prior to this extract, a scene with dark figures and mild action has been playing.

Extract 18 [CA 052109]

01 Cor: Vampire  
02 (4.0)  
03 Al:>> ((turns head to Coral)) Why you always wanna give it  
04 away:  
05 (1.9)  
06 Cor: ((turns to Al and puts hand on chin)) Well (0.1) I  
07 mean didn't you think that was a vampire gettin im  
08 ((during speech, open hand gesture to TV))  
09 (0.7)  
10 Al:>> I do:n't kno::w they look like a bunch of kids to me  
11 doin something that ain't (1.1) (have) no business  
12 doin

In line 1, Coral targets one of the dark figures who has attacked a man as a ‘vampire’. After 4 seconds (line 2), Al turns to Coral and says ‘why you always wanna give it away’, targeting her ‘giving away’ the mysteriousness in the film as problematic. After a 1.5 second gap, Coral turns to Al, putting her hand on her chin in a manner we can only describe as ‘guilty’ and accounts for having done the action that has triggered Al’s rebuke (lines 6-8). After a 0.7 second gap, Al (lines 10-12) refuses Coral’s explanation with ‘I



don't know' and further problematises Coral's turn at line 1 as well as her explanation at lines 6-8 as not even necessarily accurate, i.e. that to Al, they look like a 'bunch of kids ... doing something they have no business doing' and not vampires.

Participants may even explicitly orient to talking during television-watching as problematic even before it becomes a salient feature of the interaction. For example, Extract 19 is taken from an interaction during which Dawn and Niles are deciding which film to watch and planning how to record themselves during the film.

Extract 19 [Viney, DN2]

01 Nil:>>Mp (it) just be m(h)e: sat there an nk (.) you  
02 >>tal:king and me just i(h)gn(h)or(h)ing y(h)ou.  
03 [huhn huhn huhn ]  
04 Daw: [Y(h)ea(h)h heh ↑heh heh]  
05 (.)  
06 Nil: [°Huhn huhn huhn° ]  
07 Daw:>> [Does that annoy you.]  
08 Nil:>> .hhhhh it >doesn't< annoy me I jus: (.) I just  
09 >>don't talk during films usually.  
10 (0.2)  
11 Nil: I sort of mumble to myself.  
12 (0.2)  
13 Daw:>> Well we'll see how it goe:s and then if i:t if:  
14 (1.6)  
15 Nil: (no i that's) pretty cool.

Line 1 begins the topic, with Niles saying 'It'll just be me sat there and you talking and me just ignoring you', to which both parties laugh (lines 2-6) and Dawn responds 'yeah' (line 4) and 'does that annoy you' (line 7). Niles then says that he's not annoyed by it but

doesn't usually talk during films (lines 8-9). In line 13, Dawn begins a turn 'well we'll see how it goes and then if it ...', possibly going toward something like 'if it annoys you, we can ...'.

The above examples show the multiple ways in which talking during films can be problematised by participants. Some talking may be tolerated or even welcomed, but excessive talking or talking at inappropriate times or about inappropriate topics may be particularly problematic. Different participants have different preferences and views on talking during films, and it can be a sensitive interactional accomplishment to achieve a balance that works for all parties. Most of our data involves dyads, so resources may be used more subtly than in groups in which talkers can interact with each other despite others not talking. The strategies that non-talkers use to sanction or discourage talking may be different (more direct, perhaps) in groups.

## **[8] Conclusion**

We have examined television-watching in relation to incipient talk, sequence organisation, and turn-taking. We have discussed incipient talk and related concepts as inadequate catch-all categories and have contributed a detailed analysis of an environment sometimes described as 'incipient talk', thereby contributing to the systematic study of whether this state exists for participants. We have demonstrated that television-watching affects sequence organisation in terms of relaxed response relevance and in the complexity of responses. By engaging in different practices, participants may demonstrate to others that they are engaging different foci in a multi-focused interaction in addition to using different practices due to contextual constraints. We have also seen how turn allocation and speaker selection remain unchanged and how type-conforming and type-related responses (when responses do occur) continue to be produced despite relaxed response relevance.

Although our analysis makes significant contributions to work on possible ‘incipient talk’ situations and to audience studies, it perhaps generates more questions than answers. Whilst our dataset is rather large for a qualitative analysis, it is insufficient for robust statistical tests to fully account for differences between interactions due to small participant numbers. However, many possibilities for accounting for the differences that we have identified in this chapter lend themselves to such tests and would be useful points of departure for future research. We have described how certain constraints appear loosened during television-watching. This is not an exhaustive analysis, and further issues to examine with larger data sets (both qualitatively and quantitatively) include: viewers’ relationships to one another (ours included familiars at home, and mainly in dyads), the type of programme being watched (see Gerhardt, 2006, looking specifically at football), and the relative physical positions of participants (particularly with regard to gestures that are unlikely to be visible to the recipient as well as monitoring of the recipient by the speaker). The recent debate on ‘Mobilizing Response’ also presents an interesting challenge. Stivers and Rossano looked specifically at the likelihood of first assessments receiving responses, but our analysis highlights differences in the complexity of responses in different contexts as well. Future research may incorporate a statistical analysis of not only how turn design and action type can affect the *likelihood of a response* to a) an initiating action and b) a pursuit but also how they can affect the *complexity of a response* to a) an initiating action and b) a pursuit.

One feature of television-watching is that gestures are common forms of responses. Their frequency may be due to the fact that participants must compete with the television, if they choose to talk. However, responsive actions that are built non-vocally can also do specialised interactional work. The next chapter will explore the function of such responses during both television-watching and ordinary conversation.

## Chapter 7

### Some Uses of Gestural Responsive Actions<sup>31</sup>

---

<sup>31</sup> The version presented in this chapter was written by me. My supervisor John Rae has given substantive theoretical and editorial input, and the published version includes him as a co-author. It has been published as:

Berger, I. & Rae, J. (2012). Some uses of gestural responsive actions. *Journal of Pragmatics*, 44(13), 1821-1835.

It has also been presented as:

Berger, I. & Rae, J. (2010). *On gestural responsive actions*. 4th Conference of the International Society for Gesture Studies. Frankfurt (Oder), Germany.

Berger, I. & Rae, J. (2010, July 7). *On building responsive actions through non-vocal practices*. 3rd International Conference on Conversation Analysis. Mannheim, Germany.

**Abstract**

A central focus of research on gestures in social interaction has been their relationship to the concurrent production of talk. This chapter concerns situations in which interactants build responsive actions using gestures. We establish analytically relevant differences in approaches to the analysis of concurrent and responsive gestures before demonstrating some uses of gesture in the building of responsive actions. Previous research has established that gestures that are produced by recipients during the production of talk are often used to show stance. We show that, likewise, stance is often a salient aspect of gestures that occur following initiating actions that ordinarily make a spoken response relevant. Gestural responses may be used to do sensitive interactional work through recipients' treatment of stance as salient and by exploiting constraints that conditional relevance imposes on responsive actions. Uses of gestural responsive actions include showing stance in a sequential position, using one action type to do another, and showing that talk is forthcoming. The current findings have implications for the treatment of gesture in conversation analysis and other sequentially oriented methodologies.

**[1] Introduction**

The previous chapter addressed sequence organisation and how interactions during television-watching are different from ordinary conversation. One aspect of television-watching is that gestures are commonly used to respond to talk. However, television-watching is not the only context in which this phenomenon occurs. The current chapter examines gestural responses to talk in a variety of naturally occurring contexts.

When gestures have been examined in elicited and naturally occurring data, the central focus has often been on their relationship to the concurrent production of talk. With respect to *current speakers'* non-vocal conduct, detailed analyses of its implications for recipients' responsive conduct have been reported, for example eye-gaze (e.g., Lerner, 2003; Stivers & Rossano, 2010b) and other aspects of embodied conduct and object manipulation (e.g., Goodwin, 2007, Heath, 1986; Schegloff, 1984; Streeck, 2009b, Streeck & Hartge, 1992). With respect to *recipients'* non-vocal conduct, such matters as the display of reciprocity and stance (with respect to talk in progress) have been subjected to detailed analyses (e.g., Goodwin, 2007; Goodwin, 1980, 1998, 2006; Heath, 1986; Schegloff, 1998; Stivers, 2008). Non-vocal conduct can perform very different functions when produced concurrently with one's own talk, concurrently with another's talk, and independent of talk.

We will establish differences in disciplinary approaches to and functions of concurrent (speaker) and responsive gestures that have a bearing on the current analysis. We will then present some of the specialised work that gesture can do in the building of responsive actions. We will draw on major works in the fields of gesture studies and conversation analysis (CA) in establishing these approaches. In doing so, we will show how a multidisciplinary approach can aid in developing sound theories of gesture. Our

analysis of the uses for gestural responsive actions draws primarily on CA (See Appendix A for transcription conventions). CA has been particularly concerned with the ways in which parties in interaction use different resources to implement actions and how these actions can make relevant certain kinds of actions in response. Although CA research has considered gestures, most of the key organisations that have been identified and subjected to detailed examination within CA concern actions that are implemented through talk. (Streeck, 2009b, p. 26) Despite this, CA's focus on the sequential organisation of social interaction and how participants respond to the actions of others make it well-suited to the study of gestures in interaction.

## **[2] Concurrent Gesture**

Though by no means restricted to the analysis of gestures that occur concurrently with talk, Goodwin's research brings into prominence the close relationships between talk and gesture in the organisation of interaction. For example, one of the questions posed in his celebrated monograph is, 'How are the separate actions of speaker and hearer coordinated with each other within the turn?' (Goodwin, 1981, p. x). This is particularly significant in two respects. Firstly, with respect to the concerns of CA, it considers ways in which participants' conduct is not organised merely on a turn-by-turn basis (i.e. the largely one-at-a-time organisation of turns-at-talk) but also how speakers and recipients alike can contribute to turns in progress<sup>32</sup>. Secondly, with respect to research into non-vocal conduct more generally, particularly within psychology, it brings together the potential for a speaker's talk and their non-vocal conduct to be closely coordinated. Aspects of such coordination with regard to hand gestures have been examined in detail (e.g., Schegloff, 1984; Streeck, 1993).

---

<sup>32</sup> This is not to say that the only resources available to participants during another's talk are gestures, nor that vocal resources have been ignored (laughter is an important example). However, non-vocal behaviour comprises an important set of resources, of which Goodwin's work is particularly significant.

Although the analytical separation of vocal and non-vocal conduct is important, research under the rubric of non-vocal communication, whilst emphasising the temporal coordination of these different modalities, has tended to regard them as separate channels of communication (e.g., Mehrabian & Ferris, 1967; Scherer, Feldstein, Bond, & Rosenthal, 1985). The main analytical connection between these channels has often been an interest in their potential for conflict – and how attention to non-vocal cues might enable deception to be identified (e.g., Bond & DePaulo, 2006; Ekman & Friesen, 1969a). Although such research can be extremely useful to professions such as psychotherapy and law enforcement, it is also important to identify the *connections* between what participants do with these different modalities.

It is important to note that such connections may differ between speech communities. Aoki (2011) demonstrates that among native speakers of Japanese, speaker head nods perform at least two important functions in conjunction with other linguistic and prosodic resources: mid-turn emphasis and turn-final eliciting. Whereas emphasis can be performed in English through head movements as well (Kendon, 1972, 2002), the use of nodding to perform turn-final eliciting is a striking difference from some other languages such as English, in which gazing at the recipient, rather than nodding, is linked to completion of turns (Kendon, 1967). It is perhaps more striking, because the physical contingencies of nodding can cause the nodder's gaze to be averted, which Kendon (1967) finds in the *initial* stages of turns rather than at the ends of turns.

Whilst Goodwin's, Aoki's, and Kendon's research stresses the importance of the observable coordination of talk and gesture within the formation of speakers' conduct, the linkages between talk and gesture in terms of their psychological significance have been the subject of McNeill's research. McNeill (1985) studied gestures produced by



participants in a laboratory setting in which they watched films and upon completion narrated the stories to others who had not seen the films with them (but who perhaps had seen or heard about them previously, unbeknownst to the participant). Adults who had not been diagnosed with cognitive disabilities watched a feature length film and spent approximately one hour relating the story to listeners. Adults who had been diagnosed with cognitive disabilities and children watched films in 10 minute blocks and spent approximately the same amount of time narrating each block to listeners. He reports that few *hand* gestures were produced without accompanying talk and that all but one hand gesture were produced by narrators (McNeill, 1985, pp. 353-354; Stephens, 1983, cited in McNeill, 1985). Although the data are subjected to a thorough and insightful analysis and are carefully explicated, they are nevertheless highly constrained by the situation of their production. It is unlikely that these elicited narratives involved sensitive work or threats to face, and the speakers or listeners may not have been fully invested in the activity compared to parties involved in naturally occurring descriptions.

Although research on concurrent gestures has involved interactional as well as experimental research, it has more often addressed psycholinguistic links and evolutionary speculation than research on responsive gestures. McNeill's *So You Think Gestures Are Nonverbal?* and *Hand and Mind* are highly influential, particularly in the fields of psychology of language and psycholinguistics, and it is not surprising that such research has often focused on speakers' gestures, given McNeill's findings and theoretical focus. McNeill's elicited narrations provide a particular set of circumstances and opportunities for the use of gestures. As Kendon (1997) points out, there are many uses for gestures, and they are used in a variety of settings: 'Consequently, no simple generalization about how gestures are used in conversation is possible. ... our task is to show how the particular properties that gestures have as an expressive medium ... make them adapted for a variety

of communicative functions.’ (See also Kendon, 1985). In addition to co-occurring with a speaker’s talk, one communicative function of gesture is in responding to what a speaker has said.

### **[3] Responsive Gestures**

Looking at recordings of naturally-occurring conversation, Stivers (2008) studied recipients’ nodding during storytelling.<sup>33</sup> She found that whereas vocal continuers align with the activity of storytelling, nods do the additional tasks of claiming access to and affiliation with the teller’s stance toward the events (directly or indirectly) and indicating that the story is on track toward preferred uptake upon completion.<sup>34</sup> Note that during storytelling, unlike many other sequential contexts, an inherent obligation to produce vocal responses at particular points does not exist. Responses, whether they are gestural or vocal, tend to serve the continuation of the story rather than provide original information.

Recipients display not only hearership but also understanding, alignment, and affiliation through their gaze, facial expressions, and nodding. During descriptive sequences, a recipient’s understanding in particular is important to maintain, and Goodwin (1980) found that non-vocal practices are actually implicated during these sequences (cf. stories that often contain, but are not entirely composed of, descriptions). As with Stivers’s nods during storytelling, Goodwin observes that nods and head shakes can serve to affiliate with descriptive assessments. When a speaker does not get a desired response from the recipient, they may pursue a response (see Pomerantz, 1984b; Sacks et al., 1974). One way

---

<sup>33</sup> Listener gestures are referred to as ‘back channel’ communication by many authors (e.g., Yngve, 1970). We do not use this term, because it implies a passive role that is not necessarily oriented to by participants.

<sup>34</sup> This is at least the case for English. In Finnish, vocal continuers are used much less frequently, and over-use is considered characteristic of drunkenness (Lehtonen & Sajavaara, 1985). The primary means of showing reciprocity in Finnish is through non-vocal practices, so this environment could differ in terms of the meaning of nodding in response to stories.

to do this is to augment what has just been said through increments. Such stalling ‘might constitute an instruction to the hearer to search her own actions produced during that unit to see if there were anything absent from them that might be relevant to the speaker’s failure to advance further’ (Goodwin, 1980). Moreover, the gaze of speakers and recipients can affect the trajectory of a turn in progress (Goodwin, 1979). Interlocutors thus not only monitor the utterances of others but organise their emerging actions within turns and sequences according to the non-vocal behaviour of recipients.

A recent special issue of *Journal of Pragmatics* (Stivers, Enfield, and Levinson, 2010) involved research on question-response sequences in 10 languages. In some of these studies, non-vocal responses were not only treated *as responses*, but also as having lexical equivalents and therefore as linguistically meaningful (e.g., Rossano, 2010). They also found that in seven of the languages, responses that were composed of a gesture or involved a turn-initial gesture came significantly more quickly than responses that were composed entirely of vocal or verbal items (Stivers et al., 2009).<sup>35</sup> We take the approach of looking at responsive gestures further to identify when and why participants might produce a gestural response rather than talk in a sequential position where talk would normally be expected. In order to specify when in an interaction talk-implemented responsive actions would be expected, we must turn to the organisation of sequences in conversation.

Throughout the subsequent sections, we will draw on examples involving British and American English-speaking and Finnish-speaking adults, and all of the extracts provided

---

<sup>35</sup> Hayashi (2010) describes responsive actions in Japanese that consist of or include non-vocal components, including nods and head shakes. However, it is unclear what the positions of such components were or what meaning was attributed to each. In Aoki’s (2011) Japanese data, participants did not use *any* turn initial head nods. Although in both studies, participants were native speakers of Japanese, they may have come from different regions or immigrated to different places. Clearly, gestures do not have universal meanings, but they do not necessarily have universal placement either. It would be a worthwhile endeavour to explore the extent to which placement of gestures in responses is influenced by culture.

are of native speakers. None of the participants in the extracts have known impairments that affect their ability to engage in spoken conversation. Although one of the Finnish participants, Elisabeth, has a rather subdued conversational style from Parkinson's Disease, she nonetheless engages in spoken conversation regularly. Other settings include psychotherapy sessions (Paula, Leif, and Sally), family meals and television-watching (Coral and Al), and friends building a self-assembly chest of drawers (Sean, Katy, and Bill). Some extracts include rather lengthy silences that are not oriented to as problematic. Berger (2011/Chapter 3) and Gardner, Fitzgerald and Mushin (2009) have shown that participants orient to the local interactional environment in their orientation to silence as acceptable or as problematic. For example, as Berger (2011/Chapter 3) shows, Paula and Leif have quite different levels of silence when discussing different topics. Compare the silences in their conversations in Extract 1 on his daughter's upcoming visit and Extract 4 on computer models vs. Extract 7 on an allegory to his present situation and Extract 11, which does involve some interactional trouble in giving instructions. Silence is therefore not necessarily indicative of any interactional trouble or cognitive impairment, nor is it necessarily indicative of broad cultural differences.

### *[3.1] Sequence Organisation and Gesture*

There are many instances in conversation when affirmative or affiliative responses become relevant, and these actions can be done in a variety of ways. Consider, for example, how confirmation and agreement can be done by speech alone (Extract 1), a non-speech vocalisation alone<sup>36</sup> (Extract 2), speech with nodding (Extract 3), nodding with a vocalisation (Extract 4), and nodding with neither a vocalisation nor speech (Extract 5).

---

<sup>36</sup> Al actually produces a small nod with his Mhm, however it is not visible to Coral, who is at the kitchen sink with her back to him. We have been unable to find any cases where 'mhm' vocalisations are produced without nods, unless they are being used as a continuer during another's turn at talk.

Extract 1 [PL 121809, talk without nod]

- 01 Lei: She gonna be staying a week  
 02 Pau: Wooww  
 03 Lei: A week  
 04 Pau: That could get bloody [hh h  
 05 Lei:>> [Yeah  
 06 Pau: He [hehe hh ( )

Extract 2 [CA 062109 Dinner 3, vocalisation without nod]

- 01 Cor: =would you like me to heat up your potatuh's an green  
 02 beans for ya  
 03 (7.0 during which Al looks at plate and takes a drink)  
 04 Al: >> †Mhm

Extract 3 [CA 052109 Dinner 1, talk with nodding]

- 01 Cor: D'you still have the computer on  
 02 Al: >> Yeah ((nodding))

Extract 4 [PL 091609, nodding with vocalisation]

- 01 Lei: So you bought it about a year ago  
 02 Pau:>> Hmhm ((nodding))  
 03 Lei: You probably have Leopard<sup>37</sup>

Extract 5 [CA 052109 Dinner 1, nod without utterance]

- 01 Al: My dad had some of them  
 02 Cor: What a train set  
 03 Al: >> ((nod))

Talk is managed through sequences with beginnings and ends, and potentially pre-expansion, post-expansion, and/or insert expansion (Schegloff, 2007). *Adjacency pairs*, or

---

<sup>37</sup> A particular version of a computer operating system

*base sequences*, are such ‘beginnings’ and ‘ends’ upon which expansion can occur. Some canonical examples of adjacency pairs are question-answer (e.g., How was school today – It was okay) and greeting-greeting (e.g., Hello – Hi) sequences. *Initiating actions* (or *first pair parts*) prefer particular kinds of *responsive actions* (or *second pair parts*), and dispreferred responses are typically delayed or mitigated. Table 1 shows the action types in Extracts 1-5.

Table 1. Summary of actions types produced and implicated in Extracts 1-5.			
Extract	Initiating Action & Format	Relevant Responsive Action	Responsive Action Produced
1	Assessment as Statement	Agreement or Disagreement	Agreement
2	Offer as Yes/No Interrogative	Acceptance or Refusal	Acceptance
3	Information Seeking as Yes/No Interrogative	Dis/confirmation	Confirmation
4	Candidate Understanding as Yes/No Interrogative	Dis/confirmation	Confirmation
5	Repair with Candidate Solution	Dis/confirmation	Confirmation

Extract 1 involves an assessment and agreement, whilst Extracts 2-5 involve seeking confirmation and receiving affirmative responses. Only Extracts 2 and 3 are in the form of complete questions (‘Would you like me to ...’ and ‘Do you ...’). However, we know that

confirmation or disconfirmation is implicated through the sequential placement of ‘What a train set’ and ‘So you bought it about a year ago’ (see also Sacks, 1968 April 17/1992, pp. 752-763; Levinson 1983, p. 292). Coral’s ‘What a train set’ (Extract 5, line 2) follows Al producing a telling with a potentially unclear referent. They are watching a local news show during dinner and have been discussing a variety of topics, only some of which directly pertain to the events on screen. A story about a model train fair is on the programme at this point. Coral’s action in line 2 is to initiate repair on the phrase ‘some of them’ (line 1) and offers a candidate repair solution in doing so. This makes confirmation or disconfirmation (with correction) relevant in line 3, and Al does confirmation through a nod.

In Extract 2, Leif (line 1) produces what would appear to be a statement ‘So you bought it about a year ago’. However, his turn follows a long discussion of computer operating systems during which he has tried to ascertain which operating system Paula has. Paula has previously said that she bought her computer a couple of years ago and has listed programs she knows to have come installed on it. Leif’s action at line 1, rather than simply stating a fact, is summing up Paula’s rather convoluted answer and is seeking confirmation in order to conclude what operating system she is likely to have. Paula confirms through a vocalisation ‘hmhm’ during which she is nodding.

In our data, gestures without accompanying talk are nearly all responsive. However, it is quite possible for a gesture to do an initiating action, if the recipient’s attention has been secured. For example, we have cases of offers of food being made by, and to, adults through gestures (see also Rossano, 2011 with respect to human children and non-human primates). In institutional settings, specialised signals may be used where first pair parts might otherwise be produced vocally (e.g., a ‘go’ signal during a police or military

operation). Goffman (1967, p. 132) refers to such exchanges comprised entirely of non-vocal practices as ‘pseudo-conversations’. Seo and Koshik (2010) have also described two gestures that function as repair initiators and regularly receive verbal responses.

#### **[4] Uses for Gestural Responses**

As Kendon (2002) asks, ‘Why should we use gestures if they are just ‘words’ in another form?’ In situations where it would be undesirable to be heard (such as during police or military operations), the answer seems obvious. Even in conversation, though, people may choose to respond non-vocally to a talk-implemented initiating action that would normally receive a verbal (or at least vocal) response. In doing so, they fulfil the obligation to respond in a particular time and place (see Sacks et al., 1974) but can evade some of the constraints that impinge on vocal responses. We will consider some of the ways in which this occurs and how stance, rather than content, is often treated as the most salient aspect of gestural responses. This is not to say that gestures are used mainly to convey stance but that stance is perhaps the most accessible aspect for recipients when a gesture does not have a *precise* linguistic equivalent in a given context.

Kendon (2002) has explored uses of the lateral head shake in English and Italian. He identifies eight ways in which the head shake is used across the two languages. Although we disagree with Kendon’s conceptualisation of superlatives and intensified positive statements as ‘implied negatives’ and therefore with his statement that head shakes can always be interpreted as expressing a theme of negation, there is a clear use of the head shake as expressing negation or agreeing with something negative across most of the eight uses he describes. He identifies head shakes as being varied in length of time they take to perform, as occurring simultaneously with speech or in alternation with it, and as able to stand on their own as a ‘kinesic equivalent of a spoken particle’ or occur simultaneously



with speech as parallel or additive to the speech. It is in the sense of a 'kinesic equivalent of a spoken particle' that we are concerned with the various gestures that interactants produce. Although many of the gestures we explore do not have the fixed or specific meanings that responsive nods and head shakes have in many English dialects, they nevertheless are used in specific sequential contexts when talk would normally be expected. As we shall see, Kendon's suggestion that simple head shakes without accompanying talk serve to downplay stance or limit interactional involvement may hold true for other gestures as well.

#### *[4.1] Sequentially Positioned Displays of Stance*

Goodwin and Goodwin (2001) found that stance displays in responsive positions involved a combination of syntax, intonation, timing, and tenor. Haddington (2006) further explores the sequential and intra-turn organisation of gaze in demonstrating stance during assessment sequences. He found that both initiating actions and responsive actions contained gaze patterns that acted as resources for tracing interlocutors' stances and that interactants mirror one another's gaze in showing both positive and negative stances. In our data, no mirroring of gesture occurs when showing negative stances (but may be found in showing agreement in some cases); the possible mirroring of gaze in these circumstances is beyond the scope of this chapter. A key difference between the treatment of stance in interactional linguistic and conversation analytic literature and the examples we provide in this and subsequent sections is that the existing literature tends to address stance toward external events or third parties. However, the use of gestures alone as responsive actions, we find, is often (but not always) responsive to an initiating action that makes relevant an assessment of the speaker's qualities or competence or in which recognition or an otherwise polite response is indicated (i.e. there is a possible threat to face). Schegloff (2005) addresses the mitigated and indirect ways in which interlocutors

can raise complaints about those to whom they are speaking (often in initiating actions on the telephone, so necessarily vocally). The topical sensitivity of showing negative stance toward one's interlocutor may contribute to interlocutors choosing to respond with gesture alone in co-present situations.

Ruusuvuori and Peräkylä (2009) demonstrated that facial expressions are one way to show stance *during* another's storytelling or topic introduction. It is also possible to use facial gestures to show stance in other sequential environments. Facial expressions perform *meta-communicative* functions (Schefflen, 1974, p. 134), that is, they give information about what is being communicated. In responsive positions, they express something about what *another* party has communicated, although meta-communication is commonly thought of as something that a speaker does. Facial expressions are not merely the external expression of an internal state but have interactional significance and are used to communicate particular emotions or stances (whether actually held or not) to interlocutors (Schefflen, 1974, p. 4; Preuschoft & van Hooff, 1997). Extract 6 is an example of a negative stance being displayed through a facial gesture to express displeasure with, or that he is not convinced by, an interlocutor's reasoning. In Extract 6, Katy, Sean, and Bill have been assembling chests of drawers from instructions. They have discovered that a component has not been correctly screwed into one of the drawers and are discussing the problem and possible solutions.

Extract 6 [DIY, 'Caught Socks']

01 Kat: we:ll this one will have to have cau[ght socks then

02 Sea: [like like thi:s

03 Bil: myeaA::h

04 (0.2)

05 Sea: Well its just (.) just pull it ou[t a bit

06 Kat: [well then then it will

07           just make a hole [cuz I've already screwed through  
08 Sea:>>                                 [ ((pulls face))  
09           (0.4, Sean holds face)  
10 Kat: >It's alright it c'n stay< the:re it's not doing any  
11           harm  
12         >> (3.0, Sean motionless)  
13 Bil:     (     like tha:t)  
14 Kat:     Might not be doing's much good b't s'not doing any harm

In line 1, Katy has declared that the drawer is fine as it is and will just have to remain defective. In lines 2 and 5, Sean is showing Katy what is needed to fix it. Katy counters this, stating that she's already made a hole and doesn't want to make another one (lines 6-7). From lines 8 to 12, Sean pulls a face and maintains it. The onset of this face is following Katy's first TCU '...it will just make a hole' and at the beginning of her second TCU 'cuz I've already screwed through'. At line 9, there is a 0.4 second silence during which Sean maintains his face. Katy (lines 10-11) further justifies her desire to leave it as it is, followed by 3 seconds of silence during which Sean is motionless. Bill produces an increment (line 13) to Katy's turn at lines 10-11, and she produces further justification of her position at line 14.

Despite following the first possible completion of Katy's turn at line 6, Sean's facial expression does not explicitly target a particular problem as talk would. Even were he to say 'that' or 'it' in a negative assessment in this position, Katy's justification would be targeted as the referent by virtue of being the most recent possible referent. By adopting a non-vocal implementation, the referent is less specific. The prospect of leaving another hole, the predicament of functionality vs. aesthetics, and Katy's position not to fix it are all possible targets for Sean's facial expression. Although Katy shows that she recognises Sean's face as negatively targeting her opinion, he has been able to demonstrate this stance

without openly challenging Katy or insulting her perspective. She is able to continue giving her reasoning in a way that is tailored to the immediate context and without explicitly addressing Sean's disapproval (see also Goodwin 1979; Goodwin 1980; Goodwin & Goodwin, 1987 and cf. Kidwell, 2005).

Similarly, a potential story recipient can display reciprocity at the pre-telling stage where a go-ahead utterance would usually be expected. In Extract 7, Paula and Leif have been discussing intra-psychic processes that Leif has been experiencing, and this extract begins an apparently new topic by Leif. He has carried over hand gestures from his previous turn and used those as his turns progress, but the topical content is otherwise disconnected from his previous talk. Extract 7 has been simplified to focus on Paula's responsive actions.

Extract 7 [PL 091609 Rowan & Martin's Laugh In, simplified]

01 Lei: Rowan and Martin's Laugh-In

02 >> (3.2) ((Paula nodding))

03 Lei: Very important

04 (2.5)

05 Lei: To psychotherapy

06 Pau:>> ((hangs head and laughs quietly))

07 Lei: Very important

08 (4.8) ((Leif continuing facial and hand gestures))

09 Lei: Flip Wilson (3.8) played a number of characters ...

In line 1, Leif introduces the new topic, the late 1960s and early 1970s television comedy show *Rowan and Martin's Laugh-In*. The relevance of this is unclear at this point, and Paula (line 2) offers continuous nodding during the silence following Leif's introduction of the topic. In doing so, she demonstrates interest and the understanding that

more is on the way. Leif produces an assessment of *Rowan and Martin's Laugh-In* as 'very important' (line 3). The relevance is still unclear, and a 2.5 second silence occurs before Leif delivers the punch line of his pre-telling in the form of an increment, 'to psychotherapy' (line 5). At this point, Leif has completed his pre-telling, and Paula produces a small laugh and hangs her head (line 6). This is not only responsive to the humour of the pre-telling but serves as a go-ahead for his telling. This is separate from the nodding at the beginning of the pre-telling, which was in the service of the pre-telling itself. Whilst not entirely non-vocal due to the initial laughter (a vocalisation without a lexical equivalent), Paula's full response includes quite a strong embodied stance marker of hanging her head. The totality of this response involves Paula giving a go-ahead without talk, whilst showing a strong stance of amusement with regard to the pre-telling and the possibility of an upcoming story. (In fact, talk in this position might serve to lessen her display of amusement.) At line 7, Leif repeats the short form of his assessment (as post-expansion) and at line 9 begins his story.

#### *[4.2] Use of one action type to do another*

A talk-implemented initiating action places constraints on the kind of action that should follow and provides a framework for understanding next actions (Sacks, 1967 March 9/1992 pp. 535-546; Goffman, 1971). This creates conditions of type-relatedness and type conformity that impinge on the responsive action. Type-relatedness refers to the *type of action* accomplished by a responsive action and the extent to which it matches the actions that were implicated (e.g., acceptance/refusal to an invitation) (Schegloff, 2007). Type conformity can be defined as the property of a responsive action such that it is in the *format* that was implicated by the initiating action (e.g., yes/no to a yes/no interrogative) (Raymond, 2003; Schegloff, 2007).

One of the first conversation analytic observations concerned how a participant might do one action instead of the implicated action (Sacks, 1964/1989/1992). Sacks found that in calls to an emergency mental health centre, one of the ways in which callers could avoid giving their names was to initiate repair in the slot in which giving their names would be relevant. We are concerned with a similar phenomenon in this section. Rather than initiating a new sequence, however, interlocutors may do a different responsive action than that which was implicated. In Extract 8, Elisabeth smiles in response to an offer. Kirsi is visiting her grandmother Elisabeth and is going through a box of chocolates to find varieties that they would like to eat.

Extract 8 [Rae & Korkiakangas, 2011, Kirsi & Elisabeth, Hazelnut]

((smiles))

┌──────────┐

box...xE

01 Kirsi: hasselpähkinä  
hazelnut

box...xK

((laughter))

┌──────────┐

((smiles))

┌──────────┐

Kirsi: Elisabeth

02 (-----1-----)

Elis: Kirsi, , box -

┌──────────┐

((smiles))

┌──────────┐

((short laughter))

Kirsi has been listing varieties in the box and has located a hazelnut chocolate. She offers it non-vocally to Elisabeth by holding it out to her and smiling (line 1), followed by naming the flavour. In line 2, Elisabeth smiles responsively and both laugh. Kirsi begins a new, related sequence in line 3 'which one shall we take'. Later in the visit, we learn that Elisabeth didn't want any chocolate. Prior to this exchange, Elisabeth has been discussing matters unrelated to the chocolates. Although she has a subdued conversational style, Elisabeth is capable of speaking and does so often. By smiling, Elisabeth is able to show appreciation for the offer and continue participating in the shared activity without actually accepting the unwanted food or refusing an offer from a loved one whom she does not see regularly.

Extract 9 represents a similar situation involving showing appreciation instead of answering whether a meal was enjoyed. Coral, who has prepared their meal, has asked Al whether he liked it. Al has previously said that the steak was cooked too quickly (and was therefore tough).<sup>38</sup> Al produces a hands-together bowing gesture in response to Coral's query.

<sup>38</sup> A transcript of this prior exchange can be found in Appendix K.

03 Al: >> ((puts hands together))  
 04 Cor: [ mean it looks like it  
 05 Al: >> [ ((bows))  
 06 Al: >> ((bows again))  
 07 Cor: Well I'm glad you like it

After Al finishes his meal, Coral asks 'Did you like your dinner Al' (line 1), thereby sequentially implicating a yes/no response from him and preferring a 'yes' response. A preferred response (a positive assessment of the meal), however, would highly contradict his earlier assessment of the steak, since he has already shown strong criticism of how it was prepared. However, a negative assessment or even a mitigated assessment could be interpreted as unappreciative or at the very least would constitute a dispreferred response. Al instead responds with a gesture that involves bringing his hands together and bowing (lines 3, 5, and 6). Insofar as this gesture has a recognisable meaning, it suggests giving thanks (as in many Asian cultures) with an element of showing respect or perhaps acceptance.<sup>39</sup> Although these are positively valenced actions, as far as we know this gesture does not signify either 'yes' or 'no'. Al has thereby succeeded in producing a respectful responsive action yet has evaded explicitly answering the question put to him. This analysis is supported by Coral's response, in which she articulates a meaning that Al did not specifically produce and which constitutes the understanding of a preferred response (that Al did indeed like his dinner).

Because the hands-together bow that Al uses in Extract 9 carries less specificity than a nod, Al is able to implement a positive (but not specific) response of thanking or praise rather than an assessment of the meal. Use of a gesture in this position (for a preferred

---

<sup>39</sup> This type of bow can be done in a variety of situations and be treated as doing an appropriate action that can be understood from the context of its production: greeting, thanking, 'you're welcome', 'how do you do', and/or closing. Morris (1994, p. 194) catalogues it as greeting or thanking. Compare to words such as 'aloha' in Hawaiian, which can be understood in the context of its production to mean 'hello', 'welcome', 'love', 'peace', and/or 'good bye'.



response to an initiating action whose preferred verbal response is that of a positive assessment) is made possible by the length of time the gesture takes to implement. This thereby avoids a curt or minimal response which could be taken as incomplete and dispreferred. In summary, Al's choice of a gesture appears to make use of a weakening of the constraint of type-relatedness that is attached to talk-implemented responsive actions. He has been able to use a non-type conforming gesture where a non-type conforming utterance would perhaps not only indicate a dispreferred response but underscore that the question had not been answered. Thus Al uses a gesture to display a stance of appreciation and respect, rather than confirmation.

A further case occurs in Extract 10, which immediately follows Extract 9. Al is here listing different types of melon that he would enjoy for dessert.

Extract 10 [CA 052109 Dinner 1]

- 01 Al: Now we need some \*cantaloupe (1.1) some \*watermelon  
 02 (1.5) some \*honeyde:w  
 ((\*gestures counting beginning with right index finger,  
 middle finger, and ring finger respectively))  
 03 (3.4)  
 04 Cor: Just different kinds of fruit  
 05 Al: >> ((smiles, 'OK' sign))  
 06 (0.8, Al leaning back)  
 07 Al: °Yeah° ((turning away))

Al's suggestion of a dessert that would pair nicely with what has been served (a dessert that might conflict with this has not been prepared) appears to be a way of showing respect

for the meal (complimenting by complementing, as it were) without commenting directly on the food. Coral's response is delayed and offers a candidate understanding<sup>40</sup> of Al's list.

Our focus here is on Al's smile and ring gesture, recognisable as 'OK' and possibly synonymous with a range of positive responses from 'sure' to 'perfect'. Al appears to be engaged in moving away from the topic of the meal as prepared and to communicating a favourable opinion of the meal without stating an opinion on its quality directly. Coral's formulation of the assembly of melons that Al has proposed as 'just different kinds of fruit' (line 4) demonstrates a lack of understanding of Al's proposed ideal dessert. It appears to miss that what is being proposed is a dish comprised of specific kinds of melon. As a candidate understanding, the response that this makes relevant is confirmation or disconfirmation, with a preference for the former. Although there are grounds for disconfirmation (Al is clearly envisioning particular melons, rather than an assortment of fruit), Al does not do this. Indeed, since Al is probably engaged in seeking affiliation, disconfirmation might be highly undesirable.

He instead uses the ring gesture (or 'OK sign') as his immediate response followed by leaning back and saying 'yeah' (lines 6-7). A literal rendering of the ring gesture that Al produces here would be something akin to 'okay' or 'just fine'. As with talk, the recipient must interpret its meaning in relation to prior actions. It shows its target (Coral's candidate understanding) by virtue of its sequential position. However, if one is to render it as a complete sentence (for purposes of clarity here), the verb phrase that one might use is less obvious. It could be '...is exactly what I meant', '...would be perfect', '...would be nice too', '...is okay', among others. By using such a gesture, Al is able to convey a positively valenced response without explicitly accepting an inaccurate summary (which would

---

<sup>40</sup> This is a candidate understanding rather than simply summing up, because Coral demonstrates confusion regarding Al's prior turn. This is evident from the long gap at line 3 and tenor (not transcribed here, see video for further details) prior to, during, and following Coral's speech.

delete the meaning of his prior speech) and also avoid a dispreferred, negatively valenced response that Coral's understanding was incorrect.

*[4.3] A Possible Use: Bracketing Intervening Talk*

We present bracketing intervening talk as merely a possibility for the use of a gestural response. This is the only case we have of this kind, and it is sequentially complicated, which may account for its occurrence. It contains elements of doing sensitive work but appears to specifically avoid tangential talk that would interrupt the gesturer's ongoing spoken response to a first pair part. Further research on gesture and intervening talk or insert sequences is warranted, and we hope that this analysis of a single case might initiate such research.

In Extract 11, Paula and Leif have been discussing how Leif managed to open a computer document that she sent him. Leif has told Paula in great detail how he came to figure out the solution, and she is now asking him how he opened the file (line 1).

Extract 11 [PL 091609]

- 01 Pau: So ho::wɿ (.) did yo::uɿ (0.8) open that cz you're  
 02 right that probably is gonna happen to me  
 03 (0.9)  
 04 Pau: ((zig zag gesture)) To other folks that I send  
 05 (0.8)  
 06 Lei: Oka::  
 07 (0.6)  
 08 Pau: His documents to:  
 09 (0.9)  
 10 Pau: >What did you do<  
 11 (1.4) ((Leif has mouth open, gazing at Paula))  
 12 Pau: I know it was magic

13           (1.5)  
 14 Leif:>> ((raises eyebrows three times))  
 15           ((Paula smiles and gazes at her notes, Leif turns his  
 16           head))  
 17 Leif: Uh::m (2.8) I (1.1) saved (0.6) th' files (3.1) and  
 18           renamed them

Paula's initial question at line 1 is complicated by a self-repair on 'me' (lines 2 and 4) and a long pause in her turn at line 4 (completed on line 8) during which Leif displays understanding of her incomplete turn and begins his response (line 6). Paula's choice of words in lines 1-10 do not acknowledge that he has already told this story, and Leif displays some difficulty in formulating his response. What he actually did was quite complicated, and he has already told her. Paula continues to fully realise her first pair part, ending with 'his documents to' in line 8. At this point, it is very clear that she is requesting practical instructions that she herself can do. At line 10, Paula sums up her previous, complicated talk as 'what did you do'. This is followed by a 1.4 second gap (line 11) during which Leif has his mouth open and is gazing at her preparing to speak. Paula then initiates a side sequence (Jefferson, 1972) 'I know it was magic' (line 12), which serves to re-vivify the question and also indexes a reference that Leif has earlier made to 'black magic tricks that I learned years and years ago'. It further serves to show that she was indeed listening to his telling of how he figured out the solution the first time and that she is now asking for a simplified version for her own use.

In line 14, Leif uses a facial gesture of raising his eyebrows three times before moving on to answer her question (line 17). In lines 15 and 16, Paula smiles and turns her gaze to her notes, and Leif turns his head. Her smile and withdrawal of gaze displays orientation to Leif's gesture as adequately responsive to her turn at line 12 (rather than to her question, as it does not in itself provide an answer). Lines 12-16 thus form a side

sequence about Leif's solution as well as an insert sequence to the larger request for information and its subsequent answer. Note the lack of transition markers (e.g., 'but anyway') in line 17, which shows that Leif's response is not disrupted by the side/insert sequence. The 'um' indicates an upcoming lengthy pause consistent with the ongoing trajectory of his response.

Leif's gesture at line 14 serves to maintain his ongoing (paused) verbal response (that began in line 6) to Paula's first pair part (lines 1-10) rather than to discontinue it and contribute a vocal response to the insert sequence (begun with 'I know it was magic', line 12). If Leif had not begun talking at line 6 nor been displaying preparation to talk (especially as seen at line 11), one would not expect a verbal response to Paula's 'I know it was magic' to be potentially problematic (Sacks, 1987; Schegloff, 2007). However, it is precisely because of the expectation of contiguity (such as that Sacks and Schegloff discuss in their accounts of sequence organisation) that this side/insert sequence presents a potential difficulty. Leif's gesture responds to the initiating action of the side/insert sequence but without letting the sequence become messy. That is, he avoids producing a spoken second pair part in the midst of another, which would result in his having to later restart his current response (which he has had difficulty producing) or risk never completing it at all.

#### *[4.4] Using Gestures to Show that Talk is Forthcoming*

The non-vocal practices identified in Sections 4.1-4.3 follow, and constitute the response to, an initiating action. This is in contrast to gestures that sometimes occur in this sequential position but which are recognisable as not being *the* response. They rather serve to display that the party from whom a responsive action is due understands that a response from them is relevant and indeed may display that a response is in preparation. In this way,

an interactant can demonstrate an orientation to their obligation to respond and can signal that talk is forthcoming but not yet possible to implement.

Dickerson (2007) provides an example of how facial gestures can be produced where a response is implicated, apparently indicating that although a response has not been produced, and is not currently being produced, the implicated party is nevertheless orienting to its production. For example, in Extract 12a, following a question from a psychotherapist in a therapy session, a 14 second silence occurs (lines 6-13) before an answer is produced (line 14).

Extract 12a [Dickerson, 2007, Richard, simplified]

01 T: What did you bring (1.2) °°to the relationship:°°

02 (0.5)

03 T: What did she learn (0.5) from you:

04 (1.7)

05 T:       What did you have °°to give°

06-13 >> (14.0)

14 C: I don't think I gave her anything

However, including details of non-vocal activity, Extract 12b shows that although talk does not occur immediately after the questions (lines 1 and 3), the client is nevertheless highly active in producing a number of facial and postural movements.

Extract 12b [Dickerson, 2007, Richard]

01 T: What did you bring (1.2) to the relationship:

02 (0.5)

03 T: What did she learn [(- - - -) from you:

```
[camera cuts to client
```

04 ( - - - - - 1 - - - - - [ - - )

[blinks  
[looks down

05 T:     W[hat did you have °°to give°°  
          [looks up

06       (- [- - - - - 1 - - - - - 2 )  
          [blinks, looks down

07       (- - - - - 3 - - - - - 4 )

08       (- - - - - 5 - - [- - - - - 6 )

09                                   [purses lips

10       (- - - - [- - - - - 7 - [- - - - - 8 )  
          [eyebrow flash   [licks lips

11       (- - - - - 9 - - - - - [- 10)  
                                  [grimace

12       (- - - - - 11 - - - - - 12)

13       (- - - - - 13 - - - - - [- 14)  
                                  [Closes eyes

14 C:     I don't think I gave her [anything  
                                  [eyes open, gaze at therapist

As Dickerson (2007) points out, by producing facial gestures, the client is able to show that although he has not produced an answer to the questions that have been put to him, he is nevertheless orienting to his obligation to respond *and* to answer. Indeed, by sustaining an ongoing display of facial movements, the client is able to display that he is actively engaging in a search for an answer as well as the difficulty in doing so.

### [5] Repair on Gesture

As we have discussed in Sections 4.1 and 4.2, stance is often treated as the most salient aspect of gestural responses (as opposed to talk, where content and form tend to take primacy). Sometimes gesturers intend to communicate more specific information through their gestures, but stance is nevertheless what is interpreted by the recipient. Extract 13

involves giving a more explicit meaning of a gesture following the recipient's apparent misunderstanding of the intent. Sally has been telling Paula about friends from a different branch of her company that she met in person when on holiday to Atlanta, Georgia. Sally and Paula live in the western United States, and Georgia is in the east. The capital, Atlanta, is well known, whereas its suburbs are not. Note that Sally is usually quite mobile when sitting, because she has Parkinson's Disease. Although Sally's movements do appear at least somewhat organised in relation to her own talk, our focus is on Paula's hand gesture at lines 8-10. Sally's movements have not been transcribed here.

Extract 13 [PS 110409, simplified]

01 Pau: So they're not local:  
 02 (0.8)  
 03 Sal: No they're in it- they live 'n Atlanta  
 04 Pau: ((nod))  
 05 Sal: Well Marietta I think it is  
 06 Pau: ((two nods))  
 07 Sal: But[ uh  
 08 Pau:>> [((raises left hand, purses lips))  
 09 Pau: [HH  
 10 >> [((flops hand onto arm of chair))  
 11 Sal: Same [difference  
 12 Pau:>> [Never been don't know  
 13 (1.2)  
 14 Sal: Ah Marietta's (.) the (1.9) no:rth: (0.1) west  
 15 (0.1) corner of (0.9) >the metro area<

Paula asks a clarifying question (line 1) about the people about whom Sally has been talking, to which Sally responds in line 3. Paula receipts 'they live in Atlanta' (line 3) with a nod (line 4). Sally does a self-repair to specify that they don't actually live in Atlanta but



in the suburb of Marietta, some 6 miles outside Atlanta's ring highway. Paula's double nod in line 6 receipts this information. Sally has apparently taken Paula's nod to mean that she understands, and Sally begins to resume the prior discussion with 'but uh' (line 7). After 'but', Paula raises her left hand and purses her lips (line 8). She flops her hand onto the arm of her chair, as she produces a loud outbreath (lines 9-10). Sally's next turn 'same difference' (line 11) demonstrates that she has understood Paula's hand gesture to mean that the precise location is unimportant (a display of stance). Paula's nods at line 6 may contribute to this understanding in that the nods claim access to the details of the story. However, Paula has apparently intended her hand gesture not to be a display of stance but a display of lack of knowledge. She subsequently says 'never been don't know' (line 12).

Paula's turn at line 12 constitutes what would be termed a third position repair in talk-based analyses, targeting her hand gesture as the trouble source with Sally's 'same difference' displaying that she has understood it in a particular way. The precise location of the couple is not necessary information for the current circumstances of the discussion (Paula had asked simply whether the couple is local), and the discussion would carry on at this point unhindered. However, it is possible that the specifics of their location could come up later. Should that happen, Paula would be seen as having displayed understanding that she did not in fact have. When interlocutors are 'caught out' having pretended to know something they later can be shown not to have known, it can have detrimental effects (see Kitzinger, 2011, on pretending to remember). It is therefore advantageous for Paula to repair her gesture and be told the information now rather than in the future.

## **[6] Conclusion**

Examination of gestures that respond to, but do not co-occur with, talk indicates that gestures can accomplish sequential actions autonomously. Sometimes, these responses do

specialised work, as we have shown, particularly in sensitive sequential contexts. The specialised work that gestures can do in this position includes showing stance as a sequential action, using one action type to do another, and showing that talk is forthcoming. Analysis of a single case also suggests that a further piece of specialised work is bracketing intervening talk. In responsive positions, gestures provide interlocutors with a breadth of options to do both mundane and sensitive interactional work.

Participants such as Elisabeth who are generally more subdued in their speech may be more likely to produce gestural responses, but they do not have to by necessity. Production of a gestural response is therefore an interactional choice for these interactants as for anyone else. We do not claim that it is necessarily a *conscious* choice, but rather that one of multiple options has been selected through *some* mechanism. Possible psychological aspects of producing gestural responsive actions are a matter for further investigation. We have demonstrated that responsive gestures have interactional significance, and despite the focus in psychological fields on speakers' gestures, this is a viable area of study.

Like talk, the significance and meanings that are attributable to gestures are dependent on prior actions in their sequential environment. This is an area in which CA can be particularly useful to the study of gesture. Furthermore, although these actions occur in environments that are set up by talk, the fact that they can be used to respond evasively indicates that they have meanings of their own (to at least some degree) in the local interactional environment. That of course is not news; however, the placement of a conventionalised gesture in such an environment shows how choosing between different resources can exploit these meanings. It appears that the stance conveyed by gestures is often more salient than their content, in that the content may not have precise meanings. Interlocutors may recognise stance, even when the exact meaning of a gesture is unclear. It

is this dominance of stance in the interpretation of gestures (rather than necessarily the intent of gesturers themselves) that appears to be exploited to avoid negativity in sensitive sequential environments. That is to say that gesturers and recipients alike orient to the non-specificity of some gestures to maintain affiliation and alignment.

The use of gestures as responsive actions has implications for the treatment of gesture in conversation analysis and other sequentially oriented methodologies. Not only can gestures occupy the sequential space where talk might otherwise be used to perform a relevant action, they can do rather complex actions that could not be performed through talk (or that would be very difficult to perform through talk). Two properties of gestural responsive actions have particularly important implications for conditional relevance and preference organisation. That interlocutors can use gesture evasively means that a non-type conforming gesture can be assumed to adequately replace a type-conforming and preferred second pair part despite not having a fixed linguistic equivalent. That participants may respond with gestures as responses or preliminaries to responses when they may have difficulty producing the implicated response (or in speaking or being heard) demonstrates the presence of an obligation to respond and the availability of a range of gestures to fulfil that obligation. The capacity of such a gesture is partly a matter of how it is composed but is also a matter of the sequential position of the gesture.

## **Chapter 8**

### **Are Non-vocally Implemented Actions Subject to the Same Organisational Constraints as Talk-implemented Actions?<sup>41</sup>**

---

<sup>41</sup> The version presented in this chapter was written by me. My supervisor John Rae has given substantive theoretical and editorial input, and the version that has been submitted for publication includes him as a co-author.

### **Abstract**

In co-present interactions, talk and visible conduct are temporally and semantically inter-related. This chapter examines a methodological issue about how this inter-relationship can be examined in the context of naturally occurring social interaction. The discipline of Conversation Analysis offers a specification of how interactions are built from sequentially related actions. However, certain fundamental organisations (e.g., turn-taking and sequence organisation) that have been identified have been examined in talk. Yet talk has specific properties, in particular turn-taking, which are unlikely to be shared generally by other forms of conduct. Consequently, the role that previously identified organisations should play in the analysis of non-vocal behaviour is unclear. This chapter offers an empirically grounded contribution to this debate. We address *action formation* and *sequence organisation*, distinguish between *actions* and *turns*, and compare *response timing* for vocal and non-vocal actions. We show that although turn-taking may be talk-specific, sequence organisation can involve non-vocally implemented actions. Non-vocally implemented actions share some organisations with talk such that they can function in place of it. Non-vocal implementation also makes these actions different from talk-implemented actions.

**[1] Non-Vocal Behaviour as an Issue in the Sequential Analysis of Conversation**

Very commonly, when spoken language is used, it is in contexts where participants are visible to one another. Consequently, visible behaviour can intersect with talk in various ways. However, it is unclear just how these intersections should be examined. The previous chapter examined gestures in responsive positions, but other non-vocal behaviour occurs frequently in the context of co-present interaction as well. Early work on such behaviour in this context was primarily concerned with how people unconsciously display information to others during interactions. The identification of deception was of primary concern in the field for some time (e.g., Ekman & Friesen, 1969a).

Meanwhile, Birdwhistell (1970) was concerned with how the meaning of non-vocal practices was contextually grounded and proposed that analysis of them out of their social context would be meaningless. Although Birdwhistell is known for his contributions to the study of non-vocal behaviour through his development of the framework of kinesics (e.g., Birdwhistell, 1952, 1970), the perspective he offered on the study of social interaction more generally was much more holistic. For Birdwhistell, any understanding of social interaction had to involve the entire, complex system through which people interacted, including their intentional and unintentional sounds and movements. Although he began the framework of kinesics by adapting a structural linguistics framework to body movement, this was only a starting point and not a prescription. He remained invested in the holistic interactional context in which such movements were situated (Kendon and Sigman, 1996). Nevertheless the idea that the analysis of non-vocal behaviour could be based on a pre-existing framework that is drawn from the analysis of speech phenomena is a powerful, but problematic, idea.

On the other hand, non-vocal and vocal behaviour are clearly inter-related, and the analysis of non-vocal behaviour as a *sui generis* domain would be unsatisfactory. As Bavelas (1994) proposes, ‘one of the main functions of a speaker’s gesture is linguistic, that is, to help convey meaning to the addressee in an immediate conversational context.’ Gesture can thus transform spoken actions by working with words’ linguistic meanings to form contextually relevant actions. However, gestures and other non-vocal practices do not always occur during speech. The communicative properties of gesture are well established (e.g., Kendon, 1994), both with and without accompanying talk. We examine in the context of conversation analysis (CA), which has traditionally focused on actions formed by talk, how these practices can form complete *sequential actions* without accompanying talk.

The structure of naturally occurring interaction is a primary concern in CA (although it does not come from a structural linguistics perspective). CA is in a sense concerned with precisely the context that Birdwhistell describes, on a moment-by-moment basis. Before Harvey Sacks’s death in 1975, he and Emanuel Schegloff had begun work on the internal structural organisation of individual gestures within discourse, but this went unpublished until 2002 (Sacks & Schegloff, 2002; cf. Lerner, Zimmerman, & Kidwell, 2011 on the internal structure of tasks). Nevertheless, the seeds had been planted for the study of non-vocal behaviour from a conversation analytic perspective. Although Goodwin’s approach is primarily influenced by Goffmanian sociology and discourse analysis, his 1979 work using video recordings of naturally occurring interactions, *The Interactive Construction of a Sentence in Natural Conversation*, takes a distinctively conversation analytic tone. Since this paper was published, non-vocal behaviour has been one of many objects of study in the field of CA as well as in CA-informed research.

Although general characterisations of disciplines can be problematic, two aspects of CA that contribute to its suitability for the study of non-vocal behaviour in social interaction can be noted. Firstly, and briefly, CA adopts a data-driven orientation. Philosophically and ancestrally, this is an attitude that is shared with, if not inherited from, ethnomethodology, ‘to discover the things that persons in particular situations *do*, the methods *they* use, to create the patterned orderliness of social life’ (Garfinkel & Rawls, 2002, p. 6). More directly and prosaically, people do not do what they do in order to make theories come true, so social interaction should be studied in its own terms. Yet secondly, doing CA involves a dialectical relationship with the developing literature, which provides a set of exemplars and analytical levers for new analyses (Heritage 1984a).

However, because the organisations through which CA was initially developed were based entirely on recordings of talk (due to technological and transcription limitations), the role of non-vocal practices (including gestures, body movements, and the performance of tasks) within CA methodology has not been clearly defined despite acknowledgement of their relevance to interaction. That is to say, although CA is concerned with the moment-by-moment unfolding of naturally occurring interactions, including non-vocal behaviour, there is confusion about how to treat non-vocal behaviour on technical and theoretical levels. We argue that although turn-taking may be a talk-specific organisation, sequence organisation can involve non-vocally implemented actions in sequential positions. Although our discussion is centred within the sub-discipline of CA, we anticipate that the issues raised are of relevance to gesture scholars more generally.

Gesture-focused approaches have regarded non-vocal practices as doing particular actions or communicating particular meanings with little controversy (e.g., Kendon, 1997, p. 114; see also Kendon, 1985), and it is our aim to discuss the methodological issues



specific to conversation analysis. Rather than arguing the relative value of different methodologies in the present chapter, we aim to explore these issues through empirically based examples. For a more general discussion of gestures and other non-vocal practices in interaction, see Kendon (2004) and Streeck (2009b). Although some recent CA and CA-informed research has looked at gestures in sequential positions (e.g., Seo and Koshik, 2010; Whitehead, 2011; Berger and Rae, 2012/Chapter 7), authoritative texts make clear that sequences are constructed of *turns-at-talk* (e.g., Schegloff, 2007). If CA does not come to a unified treatment of non-vocal behaviour within sequence organisation, it may not keep up with other interactional research or may see schisms between talk-focused research and ‘multimodal’ research. As a still-young field of enquiry, such a limited theoretical scope could prove perilous to its long term viability. There is, however, a clear interest in non-vocal behaviour within mainstream CA, as the multimodal theme of the 2010 International Conference on Conversation Analysis suggests.

We aim to contribute to the debate and make a case for the inclusion of non-vocal practices in sequence organisation. We will clarify the issue of action formation in relation to turns and sequences. We argue that non-vocal practices can, even within a CA framework, be used to build actions not only in conjunction with utterances but also on their own. We will explore how non-vocal practices share some organisations with talk but not others and how these similarities and differences contribute to the construction of actions using these modalities. Detailed interactional analyses have been carried out on non-vocal practices produced by participants who have a limited capacity to express themselves through talk (as the result of developmental differences, Stribling & Rae, 2010; as an acquired condition, Goodwin 2003; or through a lack of a common native language, Streeck 1993). By contrast, the present analysis is based primarily on video data from a variety of everyday and institutional situations (e.g., household activities, school,

healthcare) involving neurotypical participants who have at least one shared language.

Most of the corpus was collected in the US or the UK and is in English, and a smaller portion in Finland and in Finnish. Jeffersonian transcription is used with transcribed non-vocal practices above lines of talk or timed silences.

Rather than pursuing a specific interactional phenomenon in detail through a single perspicuous example or through multiple cases examining the properties of specific practices, we aim to clarify methodological and theoretical points that interact in relation to non-vocal practices to develop an argument for inclusion of non-vocal practices in sequence organisation within CA. First, we address matters of *turn-taking*, *action formation* and *sequence organisation*, including the use of non-vocal practices within sequences. Second, we consider the kinds of non-vocal practices that are used in implementing actions and find it helpful to distinguish between two broad categories of non-vocal implementation: *communicative non-vocal practices* and *tasks*. Third, we compare the properties of *timing* of vocal and non-vocal actions. We then consider how non-vocal practices can be used to produce part of an action and to produce a full sequential action. We will show that in certain respects, non-vocally implemented actions share organisational features of turns-at-talk such that they can function in place of utterances. In some respects, however, non-vocal implementation makes these actions quite different to talk-implemented actions.

## **[2] Action Formation**

A key issue in the development of CA work on non-vocal practices in social interaction has been the extent to which the kinds of organisations previously identified in talk and vocal utterances might be shared by non-vocal practices. The celebrated CA work on the organization of turn-taking for conversation relates to taking turns *at talk*. Likewise,

classical treatments of adjacency pairs make clear that their parts consist of turns-at-talk (e.g., Schegloff 2007; cf. Rossano 2011). However, a single turn-at-talk can perform multiple actions that constitute a relevant second pair part plus one or more other actions (Local & Walker, 2004; Schegloff, 2007; Walker, 2010). That is to say, one turn can perform one or more actions and contain one or more turn constructional units (TCUs). When analysing interaction sequentially, it is the *action* that is described (e.g., greeting, invitation, request, acceptance, refusal), not the *turn* through which the action is performed (e.g., a sentence beginning in overlap with the prior turn, constructed of two TCUs, ending in a downward intonation, etc.). Turns are a vocal vehicle for carrying out actions; they are not what is performed but how it is performed. They occur more or less one at a time and in succession, whereas actions can occur simultaneously through the same turn (or even with the same TCU within a turn). It follows that adjacency pairs do not necessarily consist of two *turns*, but rather two *actions*.

*Action formation* refers to the process by which interactants construct actions, the ‘packages’ through which complete units of meaning are exchanged. One of the properties that Levinson (1983, p. 238) proposes actions must have in order to qualify as such is that there must be independently specifiable units onto which actions can be mapped. However, as he notes, it is impossible to specify in advance what units will perform what actions. It is the context of their production that gives units their meaning. Not having a set meaning does not automatically mean that a practice (whether verbal, vocal, or non-vocal) is not communicative nor that it does not have a sequential function or meaning by virtue of its sequential placement. The same turn can do different actions depending on its sequential position, as can a gesture or other non-vocal practice. Nofsinger (1999, p. 50) takes the questions ‘Do birds fly?’ and ‘Is the Pope Catholic?’ as an illustration of this property of talk. Although such questions in first position would display a certain ignorance about the

world and make relevant a yes/no response, the same questions in second position engender ‘yes’ and display that the initiating action has an obvious answer. Particular linguistic elements (turns-at-talk) do not necessarily have set meanings irrespective of sequential and social contexts, and it is through context that we derive deeper meaning from language. It does not stand to reason that because non-vocal practices do not necessarily have specific meanings that they cannot perform sequential actions.

Kendon (2002) has identified eight uses of the head shake in English and Italian. Let us take Extract 1 (Lerner, 2002, Extract 11 ‘Lunch’) as an example of a head shake that is used in second position.

Extract 1 [Lunch (Lerner, 2002, Extract 11/13)]

01 Ali: >Did ya have to go sign the lease for per place?<  
 02 (. )  
 03 Barb:>> ((shakes head “no”))=  
 04 Car: =Do you even have a place yet?=  
 05

In Lerner’s analysis, Barb’s head shake is treated as accomplishing an autonomous action. Alice (line 1) selects Barb to speak by virtue of Barb being the only party looking to move house. Barb, however, produces a head shake instead of speaking (line 3). Barb’s head shake falls squarely into the use of signifying ‘no’ non-vocally rather than speaking ‘no’. Her head shake is recognisable as ‘no’ due to its sequential position, i.e. following a yes/no interrogative (Raymond, 2003). Following Barb’s disconfirmation, speaker transition again becomes relevant. Here (line 4), Carol self-selects to speak and begins her turn immediately following Barb’s headshake. This is one way in which turn-taking, sequence

organisation, and non-vocal practices can intersect<sup>42</sup>. Silences in which non-vocal practices play a key communicative role can occur within pauses (intra-turn silences) as well (Collins & Marková, 1995, p. 257) and are necessary for some ritualised contexts such as certain party games (Pike, 1954/1967, p. 1).

Many of the non-vocal practices discussed in this chapter are in response to talk-implemented initiating actions. Although it is somewhat uncommon to have a responsive action implemented through non-vocal practices alone, it is perhaps even less common to find *initiating* actions implemented solely through non-vocal practices, at least in settings where talk is possible. Initiating actions provide a framework for understanding responsive actions and create interactional ‘slots’ into which responsive actions may be fitted (Pike, 1954/1967, p. 31; Sacks, 1967 March 9/1992 pp. 535-546; Goffman, 1971, 1981). Because talk is in most cases more specific in its semiotic and pragmatic content, it is able to provide such a framework more readily than non-vocal practices.<sup>43</sup> This is not to say that there are no specific non-vocal practices that can occur in first position. For example, Rossano’s (2011) ongoing research looks at human and non-human primates’ non-vocally implemented offers, and Seo and Koshik (2010) describe two gestures that perform repair initiation. Through close attention to action formation, rather than reliance on turns-at-talk, to determine the actions being performed, we will establish the importance of non-vocal behaviour for conversation analytic methodology, in particular sequence organisation.

---

<sup>42</sup> Although gestures and other non-vocal practices can affect the organisation of turns-at-talk, for example by indicating that someone plans to speak next (Streeck and Hartge, 1992) or in the management of overlap (Mondada and Oloff, 2010), we are specifically concerned with instances when non-vocal practices are used in sequential ‘slots’ (Goffman, 1971, p. 149).

<sup>43</sup> On the other hand, the wider context (or ‘interactional substrate’, Maynard and Marlaire, 1992) in which an interaction is situated provides a framework for understanding talk (and non-vocal practices, Goodwin and Goodwin, 1992); the same utterance in one context can mean something very different elsewhere or not make sense between contexts (Sacks, 1968/1995). As Kendon (1990) shows, a variety of factors including non-vocal behaviour can contribute to context.

**[3] Practices That Occur in Structurally Sequential Positions***[3.1] Terminology: Describing Behaviour*

Saville-Troike (1985) situates forms of communication along two axes, verbal/non-verbal codes and vocal/non-vocal channels (Table 1). This system is based on semiotic precision on one hand and use of the voice on the other. This is a useful starting point for the current analysis with clearly delineated categories that provide for simple and precise references to communication practices. However, let us first consider a different perspective on the verbal/non-verbal differentiation.

Table 1. Codes and Channels, adapted from Saville-Troike (1985)			
		Channel	
		Vocal	Non-vocal
Code	Verbal	Spoken language	Written language Sign languages Whistle/drum languages Morse code
	Non-verbal	Paralinguistic and prosodic features	Kinesics Proxemics Eye behaviour Pictures and cartoons

McNeill (1985) makes the (still) cutting edge statement that the designation of ‘linguistic’ as what we can write down and ‘non-linguistic’ as everything else ‘is a cultural artefact and an arbitrary limitation derived from a particular historical evolution.’ This is in itself a bold statement against ethnocentrism and ableism for his time. McNeill, in

proposing a new way to view the study of gesture, asks us to question conventional minority world linguistics and psychology of language assumptions. He goes on to provide an account of why, in his view, gesture (which he uses to specifically refer to ‘referential’ and ‘discourse-oriented’ gestures) is *verbal*. Whilst we disagree with McNeill on many of his arguments supporting this claim, he presents a revolutionary and radical approach. He argues that these gestures are verbal due to some of the relationships that he argues gestures have with talk. These proposed relationships have been critiqued, however, and new evidence has weakened the possibility of these relationships. For example, McNeill’s claim that gesture deteriorates with aphasia (but not any specific aphasia) has been critiqued by Feyereisen (1987). The elaborate use of gesture and other resources by people with aphasia (e.g., Goodwin 2003, 2004, 2011) also suggests that at least referential gestures may in fact increase with aphasia, as people try to find other ways to communicate (as others might when there are environmental barriers to using talk). Likewise, Butterworth & Hadar (1989) critique McNeill’s assumptions regarding language development, and recent cognitive neurolinguistics research has shown that people do not have uniform neural pathways for language and information (Łojek, 2009).

On these accounts, we maintain that although gestures (and other body movements) can ‘become like words’ (Kendon, 1988b), they are non-verbal (and indeed non-vocal). Because the current analysis does not deal with non-vocal verbal communication such as written language or sign languages, we use a simplified set of terms for the purpose of the current analysis. We will consider *vocal* utterances to encompass non-verbal *vocalisations* (such as ‘ahh’ or ‘mmm’) and *talk* or verbal/spoken utterances (such as ‘that’s great’ or ‘no’) (see also Mehrabian, 1972, pp. 1-2). We use ‘non-vocal practices’ to refer to body movements, task performance, gestures, and other physical actions that do not involve the voice. Because such behaviour is not simply something that occurs but rather has

interactional significance, we use the term ‘practice’ rather than the broader term ‘behaviour’ when discussing it interactionally. There are a variety of terms to refer to body movement in social interaction, which have somewhat different connotations (e.g., gesture, visible action, visible conduct, embodied conduct, bodily-visual action<sup>44</sup>). References to the visual field as a way to categorise behaviour in this sense can be problematic in that talk can also be visible. However, one demerit of the term ‘non-vocal’ is that it identifies something by virtue of what it is not. Despite this, we consider it one of the most precise terms available.

A difference between vocal and non-vocal behaviour is that the former, especially talk, is inherently communicative (Marler & Evans, 1997) whilst the latter’s communicativeness is not inherent. For example, despite occurring in direct response to a fearsome or surprising stimulus, vocalisations such as yelps and screams inform other (possibly quite distant) beings of potential danger in an area. Conversely, scratching one’s head or face ordinarily does not communicate anything but rather directly influences comfort (although these can have interactional effects, Ekman & Friesen, 1969b). Non-vocal behaviour such as this can, however, occupy a sequential position without being communicative per se (e.g., displaying uncertainty after being asked a question). This distinction can complicate conceptualisations of non-vocal ‘utterances’ (e.g., Kendon 1988b, 2004) other than those in sign languages due to the ambiguity of whether a particular behaviour is communicative and when. Needless to say, it complicates our own analysis and is perhaps one reason CA has not embraced non-vocal practices in its conceptualisation of sequence organisation. We take the position that despite it being a complex endeavour, examination of the organisation of non-vocal practices is essential.

Although there are many differences between non-vocal practices and talk, some properties

---

<sup>44</sup> Recent work by Drake and Ford (2011) has utilised ‘bodily-visual formulations’ (previously ‘bodily-visual actions’, Ford, Thompson, & Drake, 2012) to refer to practices that visually involve the body.



(such as lexicality) may not be particular to each medium (Kendon, 1988b). Whilst there are undoubtedly certain kinds of distributions relating to the opportunities for non-vocally implementing actions, talk has particular properties (e.g., constraints on turn structure and overlap) that enable specific kinds of social interaction to occur. For example, turn-taking rules can be exploited to produce long, multi-action turns. Non-vocal behaviour, likewise, has particular properties that form a very specific modality, and some of these properties will be discussed in subsequent sections.

As Kendon (1997, p. 114) eloquently states, “... our task is to show how they are so used, and to show how the particular properties that gestures have as an expressive medium ... make them adapted for a variety of communicative functions.” It is with this sentiment that we, for the purposes of the present analysis, use broad terminology for non-vocal practices to describe *function* and *sequential position* rather than form or culturally prescribed meanings. It is because we are dealing with non-vocal practices that occur without accompanying talk and that are not necessarily gestures that we refrain from making statements as to the form or meaning of such practices outside of their contextualised production. Rather, we distinguish between *communicative non-vocal practices* (which have a potentially shared communicative function) and *tasks* (which have a practical function rather than a primarily communicative function).

### [3.2] *Communicative non-vocal practices*

Ekman and Friesen (1969b) distinguish ‘communicative nonverbal behaviour’ from behaviour that is merely informative (e.g., a reactive twitch). That is, ‘communicative nonverbal behaviour’ is that which is assumed to be intentional and has a shared meaning in the context of its production. Let us take the case of nodding as a starting point (Extract 2).

Extract 2 [CA 052109]

01 Al: My dad had some of them

02 Cor: What a train set

03 Al: >> ((nod))

This practice has a widely shared and commonly understood meaning in the environment in which it was produced, namely ‘yes’. In typologies for hand gestures, gestures with such properties are commonly considered ‘emblems’ (e.g., Efron, 1941). Not all gestures that can be used in responsive actions have a specific shared meaning and therefore would not be considered emblems (cf. Ekman & Friesen, 1969b, who defines ‘emblem’ more loosely than Efron to include gestures that are understood amongst familiars as well as non-hand gestures). As Schefflen (1974, p. 37) notes, the distinction between ‘supplemental’ gestures and emblems is fluid and relative to context. The meaning of a gesture, or indeed an utterance or any act, must be deciphered by recipients; this is a continuous process during social interaction of any sort (Goffman, 1964). If a demonstrative gesture becomes known in a group of any size, it can be used instead of talk to communicate particular ideas, thus becoming emblematic. However, gestures can be used to express particular ideas without being so specific and widely known amongst members of a group. Rather, the meaning can be interpreted to an adequate extent by interlocutors. We thus refrain from placing gestures along the emblem continuum for the purposes of this chapter. In the case of non-vocal practices that we have studied, the practices observed can be categorised between two broad classes, either they involve the carrying out of task (e.g., passing an object) or they are themselves communicative in that they have appropriately specific and recognizable meanings in the particular contexts of their production. We refer to these as *tasks* and *communicative non-vocal practices* respectively.

Major typologies of gesture (e.g., Nespoulous and Lecours, 1986; Kendon, 1988b; McNeill, 1992) classify non-vocal practices according to the physical form or culturally specific meaning. Whilst Efron (1941) used ‘emblem’ (or ‘symbolic gesture’) to refer only to gestures that are not morphologically related to what they signify, Ekman and Friesen (1969b) expanded this category to include gestures that are iconically coded and gestures that are understood amongst familiars to have specific meanings. Kendon uses ‘quotable gesture’ (1990, 1992), whilst McNeill (1992) and Gullberg (1998) place gestures along continua, of which emblem is one. Despite the usefulness of classification systems, the lack of clear criteria for the categories of emblem, quotable gesture, or symbolic gesture and the fluidity of classification (Kendon, 2004, chapter 6) make this a problematic approach for the purpose of *this* analysis. We are concerned with recognisability in a fuller contextual consideration, not limited to broad cultural significance but situated use and demonstrated understanding in naturally occurring data. Therefore, we have adopted a rough classification for the purposes of communicating the present analysis, which is not intended to be a fixed classification system to supplant existing or future ones (See also Kendon, 2004, p. 107).

Non-vocal practices have varying degrees of specificity and recognisability in an individual culture, and the same non-vocal practice may have different meanings across cultures. We call non-vocal practices that can be recognised in the potentially wide *or* limited context of their production as roughly equivalent to a word or idea *communicative non-vocal practices*, in line with Ekman and Friesen’s (1969b) ‘communicative nonverbal behaviour’. As initiating or responsive actions, they can be used alone to convey meaning rather than to perform a task or emphasise co-occurring speech. A prime example of a communicative non-vocal practice is the vertical head shake, or nod, which is so widely known that it could be considered an emblem by even the most restrictive definitions (e.g.,

Efron, 1941). In many minority world cultures, a nod in response to a yes/no interrogative means ‘yes’. Another, somewhat less specific and often separately categorised (as deictic), non-vocal practice in this class is pointing in response to a ‘Where...?’ question. We will examine in later sections much less widely known practices that are quite understandable in the context of their production.

It is widely apparent that communicative non-vocal practices can be produced with or without accompanying utterances and that utterances can be produced with or without accompanying non-vocal practices. One key feature of communicative non-vocal practices that may account for their utility as sequential actions is that their meanings can be quite specific such that they can stand as *the* response or as only as a preliminary to a response, depending on the sequential context. For example, non-vocal preliminaries to responses can include tapping (Dickerson, Stribling, & Rae, 2007) or facial manipulations (Dickerson, 2007). Their meaning is also somewhat variable by virtue of not being linguistic (i.e. not talk or sign language) such that they can be used in ways that talk cannot. This issue will be addressed in more detail in Section 6 along with non-vocally implemented offers.

### *[3.3] Tasks*

Particular movements are situated in reference to the system of which they are a part, and talk’s system is language. Picking up a chess piece or reaching for a surgical instrument has its own meaning in the context of its occurrence. Non-vocal behaviours may be intercalated into the actions of individual parties and have no overt interactional significance, such as eating or knitting (Scheflen, 1974, p. 38). However, non-vocal behaviours such as these can be used to build responsive actions despite not being gestures. These practices do not serve communicative functions in the first instance but are primarily to manipulate the physical environment to achieve some interactionally relevant purpose.

Their significance depends on the context of their occurrence (cf. Goffman, 1964). One kind of non-gestural action that is implemented through non-vocal practices is that of doing something that has been requested<sup>45</sup>. Although the same task can be done without having been requested, its sequential action would be much more difficult to specify. A key difference between tasks and communicative non-vocal practices is that as responsive actions, tasks tend to be implicated by the initiating action, whereas talk is usually made relevant where communicative non-vocal practices occur. With requests that are for immediate action, the sequence is not complete simply with a verbal acceptance by the recipient (cf. requests to be fulfilled in the future) and indeed may not involve any verbal acknowledgement of the request. Request sequences involving manual object transfer are often played out with a series of thank yous from the requester during the course of the object transfer but may not involve any vocalisations on the part of the request recipient (Rae and Guimarães, 2010). Extract 3 is one such an example. Here Katy produces a request for a tool to be passed to her (lines 1-2), and subsequently Sean does just that (lines 5-7). The sequence ends at line 8 with Katy producing a final ‘Thhanks’.

Extract 3 [DIY, Rae & Guimarães, 2010]

01 Kat: .hhh Would you be so kind as to pass me over thuh  
 02 hammer  
       ((Sean searches for hammer))  
 03 (0.2)  
 04 Kat: °Thankyou°  
       >> ((Sean picks up and passes hammer))  
 05 (2.0)  
       >> ((Katy takes hold of hammer))  
 06 Kat: Cheers,

<sup>45</sup> Even doing a gesture in this position would not involve doing the gesture in a communicative function. Although a request to, for example, repeat a gesture could be in the form of a request, it would most likely sequentially perform the action of repair.

```
>> ((Sean releases hammer))
07      (0.3)
08 Kat:  Thhanks.
```

If it is not verbal acceptance of the request that is required to move the sequence toward completion, at what point is the second pair part considered adequate? As Extract 4 shows, it is not even the beginning of the task but the completion of the task that completes the second pair part. The entire task occupies the sequential slot in response to a request for immediate action. Whilst watching a film and sitting next to Coral on the sofa, Al makes a request for a pillow, the fulfilment of which is oriented to and begun but is not completed until much later. The participants orient to the request being incomplete and do not close the request sequence until the transfer has been fully accomplished. This request sequence happens concurrently with a discussion of the events on screen, contributing to the lengthy time it takes to complete the request. Coral treats the events on screen as the priority, whilst Al treats his request as the priority.

Extract 4 [CA 052309]

```
01 Al: >>Get the pil[l- get the pillow
02 TV:      [((brakes screeching on screen))
03      (1.6) ((Coral gazing at TV))
04 Cor:  ((turns to Al, wide-eyed expression)) Whoa
05      (4.9)
06 Al: >>Get the pillow
07 Cor:>> ((reaches for pillow))
08 Cor:  O:kay
09      (1.9)
10 Cor:  He just let one a his secrets out
11      >> ((keeps pillow held opposite side from Al))
12      (1.7)
```

13 Cor: You saw what he did didn't you  
 14 (1.4)  
 15 Al: He stopped the car  
 16 Cor: Yes: (0.5)((raises arm to waist height)) with his \*arm  
 17 ((\*uses arm to gesture emphatically))  
 18 Al: Things happen to ( ) miraculous[l-  
 19 Cor: [Ih- [It dented  
 20 Al: [I  
 21 Cor: Look how it dented  
 22 Al: I: I got that  
 23 Al: >> #GIVE ME THE PILLO'#  
 24 Cor:>> ((Gives Al the pillow))  
 25 Al: ((puts pillow beside him, behind Coral,  
 26 Coral leaning forward))  
 27 Al: KAY  
 28 Cor: ((leans back))  
 29 (2.9)  
 30 Cor: His strength ((looking at Al then turns back to TV))

Al requests a pillow (line 1) that is on the other side of Coral. His initial try occurs during a dramatic moment on screen, and the sound of screeching brakes overlaps 'pillow'. He reissues his request in the clear. Coral is evidently engrossed in the film, and her 'whoa' (line 4) is responsive to the events on screen. Al again reissues his request in line 6, and Coral reaches for the pillow (line 7). Her 'okay' is continuing her commentary on the on screen events, as is her talk at line 10 'he just let one of his secrets out' (that the character is not human). She has been holding the pillow on her side opposite that of Al and continues commenting on the film, seeking Al's input at line 13 ('you saw what he did didn't you').

Al does not embrace the amazement that Coral is displaying and responds simply 'he stopped the car'. Coral (line 16) states why it is amazing, which Al dismisses as miraculous things just happening sometimes (rather than necessarily everyone now knowing that the character is a vampire). Coral reiterates her amazement with 'it dented' (line 19) and 'look how it dented' (line 21), and Al again demonstrates that this is not of interest or news to him, 'I got that' (lines 20 & 22). Coral is still holding the pillow on the side away from Al and therefore has not finished fulfilling the request. She has begun it, and Al has not reissued the request again up until this point. At line 23, Al in a loud and screechy voice says 'give me the pillow', and Coral finally gives Al the pillow (line 24). After adjusting his position (line 25-26), Al shows that the request has been fulfilled to his satisfaction with 'kay' (line 27), closing the sequence. Coral continues her commentary on the events on screen at line 30.

Extracts 3 and 4 show that 1) responsive tasks can be completed without accompanying utterances and 2) the sequence is not treated as complete until the responsive task has been completed. It is not only through strictly communicative practices that social life is developed but through engagement with one's environment. This can take on a sequential structure such that performance of tasks can be made relevant by a initiating actions and subsequently treated as adequately or inadequately fulfilling the implicated action.

#### **[4] Action Timing: Onset, Duration and Transition**

##### *[4.1] Action Onset*

A fundamental insight of early, foundational work in CA, and a central concern of subsequent analyses, is that social interaction involves series of actions that stand in specific relationships to one another. A turn-at-talk can implement one or more actions



(e.g., requesting information) that makes relevant, or perhaps even implicates, a specific kind of action in response. When a response does not occur under these circumstances, it is noticeably absent, with the silence attributable to particular parties (Sacks, Schegloff, and Jefferson, 1974). Related to such considerations are matters of timing, namely when the relevant or implicated action should occur. Routinely, such a responsive action should occur as soon as the initiating action is complete. For example, in Extract 5, Al produces a turn indicating that a particular bus driver is 'not on route right now'. Coral's request for clarification of the target 'right now' (line 2) implicates on its completion that a) Al should produce an answer and b) the answer should be produced now. He does both, and the conversation moves forward, with Coral producing a possible account for the bus driver's absence.

Extract 5 [CA 052109]

- 01 Al:     He's not on route right now  
02 Cor:>> You mean today  
03 Al: >> Yeah  
04 Cor:    Maybe he took off for the day

Inevitably, a number of issues can come into play in the production (or lack thereof) of a responsive action, and cases where responsive actions do not occur (or are delayed) have been subjected to detailed research (e.g., Pomerantz, 1984b). Whilst Extract 5 involves both initiating and responsive actions implemented through talk, some talk-implemented initiating actions implicate the performance of a task, a particular non-vocal practice. Whilst the parties to whom such requests are made may produce talk prior to or alongside the implicated passing, occasionally only the task is performed (with no accompanying utterances). An example is shown in Extract 6, a reproduction of Extract 3.

Here Katy produces a request for a tool to be passed to her (lines 1-2), and Sean passes the tool.

Extract 6 [DIY, Rae & Guimarães, 2010]

01 Kat: .hhh Would you be so kind as to pass me over thuh  
 02 hammer  
 >> ((Sean searches for hammer))  
 03 >> (0.2)  
 04 Kat: °Thankyou°  
 ((Sean picks up and passes hammer))  
 05 (2.0)  
 ((Katy takes hold of hammer))  
 06 Kat: Cheers  
 ((Sean releases hammer))  
 07 (0.3)  
 08 Kat: Thhanks.

Whilst there appears to be some latency in Sean performing the task (line 3), his engagement with locating the requested object does in fact immediately follow Katy's request. It should be noted that the physical contingencies of the setting are such that the overall action of complying with the request has its own phase structure. In order to carry out the passing of the hammer that has been requested, Sean must, as part of this action, first locate and pick up the hammer. The initiating action here implicates a non-vocal response and gets a non-vocal response. Communicative non-vocal practices likewise routinely start at the place where an action implemented though talk might start, as in Extract 7.

Extract 7 [CA 052309]

01 Cor: Did she say Charlie  
 02 Al: >> ((nod))  
 03 Cor: I thought that was her dad

In Extract 7, Al's nod (line 2) responds positively to Coral's yes/no interrogative in line 1. Al's choice of a non-vocal practice, rather than talk, apparently relates to the context. He and Coral are watching television, and the choice of modality appears to demonstrate (along with sustained gaze at the television) that the television, rather than interaction with Coral, is Al's primary project of the moment (Viney and Berger, 2010; Berger and Viney, 2010). Al's nod has the same sequential and relative temporal placement as his 'yeah' in Extract 3.

Non-vocally implemented responses are subject to similar onset timing constraints as talk, in that some kind of action is relevant at a transition relevance place<sup>46</sup> (TRP). Such an action may be comprised of talk, non-verbal vocalisations, and/or non-vocal practices; however lack of any of these can occasion a pursuit by the producer of the initiating action. By beginning a non-vocally implemented response just when a talk-implemented response would be relevant, respondents can delay, or even avoid, the production of a talk-implemented dispreferred response and instead produce (or imply) a preferred response. In Extract 8, Al gives a bow in response to Coral's query as to whether he enjoyed his meal. His gesture begins only slightly before Coral's pursuit (following a 1.4 second gap, during which Al is chewing) and is in overlap with it.

---

<sup>46</sup> Although 'transition relevance place' is usually used to refer to a point where transition between *speakers* or *turns-at-talk* becomes relevant, we use it here to refer to a point where transition between *participants* or *actions* becomes relevant. At the point in which the TRP begins, there is not necessarily any indication which modalities will be used. By applying TRP only to instances where talk is produced as the responsive action, researchers would be engaging constantly in analysis of actions yet to occur and to which participants themselves would not yet have access.

Extract 8 [CA 052109]

01 Cor: Did you like your dinner Al  
 02 (1.4) ((Al chewing))  
 03 Al: >> ((puts hands together))=  
 04 Cor: I[ mean it looks like it  
 05 Al: >> =[(bows)]

Simultaneous starts (Jefferson, 1984b) in talk have been described, and sequentially such simultaneous starts involving non-vocal practices appear to function similarly to simultaneous starts involving two speakers. Extract 8 is an instance of a responsive action being performed entirely non-vocally. Whether a non-vocal practice is followed by utterances, the timing of a non-vocal response at the point at which a talk-implemented response would be expected demonstrates responsiveness. By beginning a non-vocally implemented response just when a talk-implemented response would be relevant, respondents can delay, or even avoid, the production of a talk-implemented dispreferred response and instead produce (or imply) a preferred response.

#### *[4.2] Permeability of Non-vocal Practices*

Talk commonly happens with one speaker at a time<sup>47</sup>, probably due to limitations in processing speech from multiple sources at once (Stifelman, 1994; Mousavi, Low, and Sweller, 1995). Overlap is most common at the ends and beginning of turns due to the projectability of talk through a combination of syntax and prosody (Jefferson, 1984b, 1986, 2004; Couper-Kuhlen and Ford, 2004), and Lerner (1996) has described the semi-permeability of grammatical units in turns-at-talk beyond terminal overlap. In all of these

---

<sup>47</sup> It is a common misconception that Sacks et al. (1974) observed that transitions between speakers commonly happen without overlap. However, a careful reading of this influential paper reveals that in fact they were referring to talk more generally (i.e. that during the majority of any given stretch of time, people do not speak at the same time). On the other hand, 'overlap is common but brief'. This refers to mid-turn overlap as well as terminal overlap and simultaneous starts. Although this is not necessarily how the majority of transitions occur, overlap during transition or at *possible* completion is common but brief.

cases, overlap refers to talk overlapping talk. Non-vocal practices, however, neither need to occur one-at-a-time nor are necessarily projectable. Tasks such as object passing have been shown to have clear trajectories that are oriented to by participants (Rae & Guimarães, 2010; Lerner & Raymond, 2008), as do many individual gestures (e.g., Sacks and Schegloff, 2002). Because gestures can be understood even during speech, such projectability is unnecessary for communicative function. Their ability to co-exist with talk and other non-vocal practices (provided recipients can see them) allows interlocutors to produce sequentially oriented actions that are not subject to turn-taking. Goodwin (1980) and Stivers (2008) have shown how non-vocal practices can perform a variety of interactional functions within the space of another's turn at talk, and Kendon (2002, p. 148) also alludes to this property of head shakes more specifically.

Although the onsets of non-vocally implemented responsive actions appear to be subject to the same temporal considerations that relate to talk-implemented responsive actions, the endings of these actions, when performed as a series, commonly appear to be less determinate than talk. Take Extract 9 for example.

Extract 9 [CA 052109]

01 Cor: Oh really=outdoor wedding?

02 Al: >> ((nods several times))

03 Cor:>> O]h that's awesome

Al responds to Coral's question (line 1) by nodding (line 2); this nodding only ceases once his recipient starts a turn-at-talk (line 3). Often non-vocal practices do not spontaneously come to completion but rather continue until a subsequent turn by a recipient begins. We are not aware of any research that has been able to uncover an overall structure of such series of repetitive gestures. Although individual practices may have projectable ends,

series of them are not known to have an overarching structure that informs interlocutors as to how many, for example nods, will occur or how long the series will last.

### **[5] When Non-Vocal Practices Constitute a Part of the Overall Action**

In the previous section, we dealt with cases in which there was an initiating action that makes relevant a talk-implemented response but which is immediately followed not by the implicated speech but by a non-vocal practice. In the following cases, however, the non-vocal practice does not accomplish the implicated response but rather appears to be preliminary to it. For detailed analyses of a 14 second silence during which an interactant shows that an answer to a question is in the works, see Dickerson (2007) and Berger and Rae (2012/Chapter 7). Whilst the activities examined here occur in a specific environment (following an initiating action in which a response is implicated), they can be compared with other cases where non-vocal activity is used to display attention or disattention. Goodwin (1981) has shown how engagement with an activity can account for temporary disengagement from conversation. Just as non-vocal practices can be used to account for temporary disengagement from talk, they may be used to delay talk in more direct ways. They may be used in a similar way to perturbations and yes/no at the beginnings of responses (Schegloff, 2007) to delay a dispreferred response. In Extract 10, following a long gap and pursuit, Al gives a horizontal head shake prior to criticising how Coral cooked his steak. The position of this headshake is that of pre-disagreement (Schegloff, 2007, p. 72).

Extract 10 [CA 052109]

01 Cor: Was your steak tender enough?

02 (2.3)

03 Cor: huh?

04 Al: >> ((single headshake)) >Cooked it too fast<

05 (0.8)

06 Cor: I cooked it too fa:st

In Extract 11 (an extension of Extract 10), Al again uses an initial non-vocal practice followed by criticism (line 8) in response to Coral's repair initiation (line 6). In this instance, the nod serves to confirm that Coral's understanding of line 4 is correct, whilst slightly delaying the spoken dispreferred element. The delay serves to mitigate the dispreferred response and lessen its negativity (i.e. the response is neither abrupt nor displays eagerness to respond negatively).

Extract 11 [CA 052109]

04 Al: ((single headshake)) >Cooked it too fast<

05 (0.8)

06 Cor:>> I cooked it too fa:st

07 (0.7)

08 Al: >> ((nodding)) >Y'cooked it too fast<

09 (0.8)

10 Cor: Really:

There is a long tradition of research that looks at the ways in which gestures can augment talk-implemented actions, including an underlying organisation of gestures within sequential environments. Within-utterance gestures have been described by Heath (1992) in terms of how they aid context and understanding in relation to the ongoing talk and setting. Such gestures may be mimicked by subsequent speakers, at which point they are termed *return gestures*. Although these gestures are not necessary to convey the meaning of the actions, they help to structure responses within the ongoing sequence. In doing so, they function to establish a shared display of intersubjectivity. (de Fornel, 1992; Koschmann and LeBaron, 2002, p. 262) Such sensitive interrelation is not limited to co-

occurring vocal and non-vocal practices but can involve them in alternation with each other. Olsher (2004) describes how talk-based actions can be completed by gestures following a partial turn-at-talk. In such instances, the end of the turn is projected but rather than complete the action through talk, the speaker uses a gesture. Recipients do not begin their responses until the gesture is completed. Olsher analyses this phenomenon as an incomplete turn plus a gesture. However, we would argue that this is a case in which gestures (and perhaps other non-vocal practices) may come closer to turn status. Whilst turns consist of talk under normal circumstances, gestures in this position not only constitute part of the action but part of the turn as well. This is made possible because an end is projected with talk but completed through gesture, and recipients treat it as part of the turn by not speaking until it is complete.

#### **[6] When Non-Vocal Practices Constitute *the* Action**

So far we have been concerned with the temporal organisation and placement of non-vocal practices. This has included their placement within turns-at-talk, their onset following initiating actions, and their duration in relation to talk. We now turn to non-vocal practices used in lieu of talk in sequential ‘slots’. Whitehead (2011) has described three distinct forms of nods that can perform distinct functions in third position with or without accompanying talk. In these instances, the speaker of a first pair part acknowledges receipt of the second pair part with a sequence closing third. Whitehead shows that the three types of nods he has described mark the second pair part as news, not news, or as specifically dispreferred news. The deceptively simple nod is used to do particular work in specific sequential environments. In this case, it is used in third position, but it and others regularly appear in initiating and responsive actions. We will now explore some of the uses of non-vocal practices that are used in lieu of talk in these positions.



*[6.1] As Responsive Actions*

An initial and obvious observation that nonetheless merits articulation is that the familiar routine relationship of contiguity that is found with talk-implemented responsive actions also holds for non-vocally implemented responsive actions. The contiguity stemming from the creation of ‘slots’ makes possible the use of quite non-specific gestures in responding to talk. This contiguity can be seen quite clearly in Extract 12, in which Al is actively eating a large bite of salad yet sustains engagement in the conversation without producing any utterances.

Extract 12 [CA 062109]

- 01 Al:     [ ((taking a bite of salad))  
 02 Cor:   [ ((turning to Al)) She fell out; ((gaze remains on Al))  
 03 Al: >> [ ((bite turns into nodding))  
 04 Cor:   [ (   ) she's pregnant too isn't she  
 05 Cor:   HHhhh!  
 06 Cor:   He' going to find out she's pregnant  
 07 Al: >> ((smiling chin-thrust))  
 08 Cor:   ((turns back to television, takes drink))  
 09 Cor:   ((turns to Al)) mmhmm  
 10 Al: >> ((turns head to Coral, smiles))  
 11 Cor:   ((turns to face forward))  
 12 Al: >> ((single nod and turns head forward))  
 13 Cor:   ((gets up and walks toward kitchen)) yep fraid so

Al has been taking a bite of salad, has his chin lifted, and is holding a forkful of food.

Coral's turn (line 2) begins as Al is lowering his chin from this position and is still manoeuvring a piece of salad into his mouth. She produces a candidate understanding about the events on screen with rising intonation whilst turned to face Al. Al lowers his chin, and the rising from this position forms the first stroke of a series of small nods (line

3). Whilst it would be physiologically possible for Al to produce speech or at least vocalisations, this would be difficult. Selecting a non-vocally implemented responsive action allows him to respond straight away whilst continuing his work on the salad. Similar cases occur in lines 7, 10, and 12. In line 6, Coral gives a prediction of what is going to happen to the characters. Al responds straight away with a smile and chin thrust (reminiscent of laughter but without vocalisation), and is thereby able to respond and continue chewing his food (line 7). By producing these facial and upper body movements following Coral's initiating actions, his actions fit within the slots that Coral has created. She indeed treats his movements as appropriate responsive actions.

Non-lexical items can be used to communicate when there is a problem with speaking or hearing, even when parties do not have organic difficulties in these areas. These practices may be followed by lexical items, when the problem has passed. For example, Wiggins (2002) shows how during chewing, people may produce 'mmm's with to positively assess the flavour of the food they are currently eating. This is often (but not always) followed by a verbal assessment of the food once it has been swallowed. Gardner (2002), conversely, notes not fewer than five kinds of 'mmm' with different communicative functions. Such use of alternative forms of communication can also include non-vocal resources. In Extract 12, Al produces a series of non-vocal responses to Coral's talk, thereby non-vocally maintaining engagement in the conversation. During this extract, Al is chewing and watching television, and by producing non-vocal responses, he is able to continue eating, not talk with his mouth full, not compete with the television to be heard, and not contribute to the disruption of his hearing the television.

People regularly use non-vocal practices during television-watching in instances when *speaking* is not a problem, but rather *being heard* by a recipient or *hearing* what is on

screen is the potential difficulty. This is especially true of a variety of nods and headshakes in response to evaluations, assessments, and yes/no interrogatives. (Viney & Berger, 2010; Berger & Viney, 2010) With regard to problems with speaking that are not situational, a rich body of research has developed involving people diagnosed with aphasia or autism. This research has revealed how non-vocal practices may be used to augment or replace lexical communication and are often performed in conjunction with non-lexical vocalisations to provide further clues to complex meanings (e.g., Goodwin, 2003, 2004).

Goodwin and Goodwin (2001) demonstrate that stance can be shown during responsive talk through prosody. Goodwin (1979) shows that turns at talk can change in their course through the non-vocal interaction with interlocutors. Goodwin (1980) takes this further to show that speakers and recipients engage in continuous monitoring and performance of stance and adjust their actions accordingly. Berger and Rae (2012/Chapter 7) show that stance can be communicated as a responsive action built through gesture and that such a response can serve to avoid overt conflict in cases of negative stance. They demonstrate how gestural responses can be used to do sensitive interactional work where vocal responses with similar meanings would be conspicuous as not performing the implicated action.

Since non-vocal practices often take more time to implement than talk with similar meanings, interlocutors may be able to implement actions that are not possible through vocalisations or speech. By producing a non-vocal practice alone, a participant can avoid a dispreferred response or a disingenuous preferred response yet fulfil the obligation to respond. Take for example Extract 10, above, in which Al produces a non-vocal response to a yes/no interrogative by Coral ('Did you like your dinner Al'). Coral has previously (Extracts 11 and 12) asked Al whether his steak was tender enough, and received a

dispreferred response. Al has exhibited unwillingness to elaborate on the criticism or to absolve Coral of responsibility for the steak's toughness through his lack of uptake following multiple attempts to elicit agreement by Coral.

Responsive actions can be built using non-vocal practices that are ritualised in the context of their production. For example, in Igbo marriage proposals, the girl/woman to whom a boy/man has proposed is expected to turn shyly and run away to rejoice if she accepts or to stand silently if she refuses (Nwoye, 1985). In either case, the proposed-to girl/woman is expected to be so emotional that she cannot speak, despite the central role of spoken exchanges in Igbo culture. The central role of talk is evident in the status of being people who do not speak to one another. Not speaking to someone means that they are one's enemy, and when approaching them a gesture signalling to others that the person is bad and not someone to whom one speaks is utilised. Whereas in Euro-American English, non-vocal implementation often serves to downplay meaning or make subtle points, the communicative function of non-vocal implementation in societies such as the Igbo carries strong meaning in contrast to talk due to the central role that spoken exchanges play in Igbo culture.

#### *[6.2] As Initiating Actions*

In previous sections, we have been primarily concerned with non-vocal practices that occur as responsive actions (or second pair parts). We have primarily looked at responsive actions, because their timing is made visible to us as analysts as well as to participants, whereas initiating actions can generally happen at any time without parties being accountable for their timing. However, it is possible for non-vocal practices to occur as initiating actions, if the recipient's attention is already secured. An obvious and common use of non-vocal practices in first position in many cultures throughout the world is that of

greeting, often a hand (e.g., wave) or facial gesture (e.g., nod). In institutional settings, specialised signals may be used where initiating actions might otherwise be produced vocally. In some instances, such as a ‘go’ signal during a police or military operation<sup>48</sup>, producing a vocal initiating action could endanger interactants. Although producing an utterance is physically possible, it is to the producer’s advantage to use non-vocal means.

In other situations, talk is not necessary to accomplish the action, and the producer of the initiating action simply dispenses with it in favour of producing an entirely non-vocally implemented action. Extract 13 represents one such an instance, where in fact the entire sequence, including both the object transfer sequences (lines 2-3 and 5-6) and the overall sequence of Kirsi helping Elisabeth (lines 2-6), is carried out non-vocally. Elisabeth has been telling a story, and the spoken turn in line 1 is in the service of continuing that story. Because it contains a repair and ends with an incomplete TCU, we have not provided a gloss in the translation.

Extract 13 [Korkiakangas, Kirsi & Elisabeth]

01 Eli: ne kävi panemassa (2.8) kävivät (1.1) kattomassa onko siellä  
*they were putting (2.8) were (1.1) looking whether there is*

02 Kir: ((reaches for sweet))

03 Eli: ((moves sweet toward Kirsi's hand slightly))

04 Kir: ((takes and unwraps sweet in same position))

05 Kir: ((moves unwrapped sweet slightly closer to Elisabeth))

06 Eli: ((takes sweet and eats it))

<sup>48</sup> Analysis of data of this sort may prove beneficial to the understanding of the role of non-vocal practices in initiating actions. Filming actual police or military operations could prove dangerous in and of itself and/or a risk to the security of the unit or larger institution. However, conversation analysis of training exercises could be helpful not only with regard to basic/pure research but also in recognising difficulties within units. For example, Nevile (2006) demonstrates through the analysis of successful commercial passenger flights and a fatal crash due to crew error how more than very rare occurrences of overlap in cockpit talk can indicate that crew members are not functioning as a cohesive and cooperative unit.

In line 2, Kirsi non-vocally offers to unwrap a sweet for Elisabeth, and Elisabeth accepts non-vocally (line 3). Kirsi unwraps the sweet in line 4, and offers the unwrapped sweet to Elisabeth (line 5). Elisabeth then takes the sweet and eats it (line 6). Goffman (1967) referred to exchanges composed entirely of non-vocal practices as ‘pseudo-conversations’, but non-vocal initiating actions may also receive vocal responsive actions. Extract 14 is one such an example. Here, Kirsi offers Elisabeth chocolates non-vocally (line 1), and Elisabeth responds that she doesn’t want any more (line 2).

Extract 14 [Korkiakangas, Kirsi & Elisabeth]

01 Kir: ((holds open box of chocolates out to Elisabeth))

02 Eli: emm [ä:ota enää

umm [I’m not taking more

umm [I’m not taking any more

((closes and withdraws box))

03 Kir: [etsä: ota enää aha

[You’re not taking more okay

[You’re not taking any more okay

Kendon (1988a) relates Goffman’s comments in *Interaction Ritual* (1967, p. 2) that the proper study of interaction entails the study of ‘syntactical relations among the acts of different persons mutually present to one another’ to the ideas of early gesture researchers such as Birdwhistell and Scheflen. This is a similar approach to Pike (1954) as well, who argues that in order to fully understand co-present interactions, non-vocal behaviour and talk must both be analysed with regard to their specific placement in the interaction, whether at the action level or within smaller units. If one were to analyse this sequence using only vocal information, it would appear to be an announcement by Elisabeth that she is done with something, which Kirsi then simply repeats. However, it is in fact an offer by

Kirsi (line 1) to which Elisabeth is responding that she is done with the chocolates (line 2).

Kirsi then responds to Elisabeth's refusal by closing and withdrawing the box of chocolates and verbally acknowledging her refusal (line 3).

Non-vocal practices in first position can be augmented with vocal practices just as in responsive actions (see Berger & Rae, 2012/Chapter 7). For example, Extract 15 includes a repair initiation accomplished by Paula (lines 4-6), with a loud outbreath (line 5) as the only audible component.

Extract 15 [PS 110409]

01 Sal: Well Marietta I think it is

02 Pau: ((two nods))

03 Sal: But[ uh

04 Pau:>> [((raises left hand, purses lips))

05 Pau:>> [HH

06 >> [((flops hand onto arm of chair))

07 Sal: Same [difference

Non-vocally implemented initiating actions are rare even compared to non-vocally implemented responsive actions. The relative rarity of non-vocally implemented initiating actions may be because talk-implemented initiating actions provide a framework for understanding responsive actions (whether vocal or non-vocal). Highly ritualised exchanges such as greeting and goodbye gestures have their meaning from the context of seeing someone anew or having closed the interaction as well as their ritual significance. However, most non-vocal practices do not carry this degree of specificity. Since many non-vocal practices do not have exact linguistic meanings (aside from sign languages), even when context is provided, it is very difficult for them to provide the kind of framework that an initiating action must provide in order to receive a meaningful response.

The prevalence of offers as non-vocal initiating actions may be due to the context that can be provided even in the absence of talk. A non-vocal offer characteristically involves holding the arm away from the body in a particular manner, with the object being offered within reach of the recipient (Kidwell, 2010; Kidwell & Zimmerman, 2007, p. 605; see also Rossano, 2011). The context is provided by the presentation of the object with which the recipient is to engage; an empty hand in this position might cause understanding difficulties.

## **[7] Conclusion**

McNeill raised the possibility that talk and gesture are more intimately related than had been previously thought, specifically that gestures and speech are intimately and linearly related on a neurological and developmental basis. However, his studies only concern laboratory situations and focus on proposed underlying psychological processes rather than the interactional situations in which gestures occur in day-to-day life. Although talk and non-vocal practices are not related in all of the ways that McNeill proposed, particular aspects of non-vocal practices partake of the same organisation as talk. The onsets of non-vocally implemented responsive actions appear to be governed by the same kinds of considerations as those governing those that are implemented through talk. Conversely, however, although the onsets share this feature, the endings of non-vocally implemented actions can differ from talk in some respects. Some non-vocal practices, such as Al's bow, are fairly finite, whereas others, such as nodding, are less constrained. This undoubtedly relates to the internal structure of non-vocal practices, or lack thereof in series of repeated gestures, and the particular property of projectability that talk's TCUs have (Sacks et al. 1974).



In responding to a talk-implemented initiating action, participants sometimes produce responsive actions that are entirely non-vocal. Whilst some initiating actions specifically make a non-vocal practice relevant, interlocutors will on occasion respond with a non-vocal practice although talk has been made relevant. In comparison with talk-implemented actions, such actions appear to have two particular features:

- (a) The type of action that they implement can be less specific; their type-relatedness and type conformity to initiating actions appear to be weaker.
- (b) In certain respects, they are less constrained by turn-taking considerations.

These features appear to be relevant, at least on some occasions, for participants adopting them rather than talk. In some respects then, non-vocally implemented actions are like talk-implemented ones, yet in other respects they partake of a different organization. Whilst we have focused on the methodological implications for CA, the detailed account of sequence organisation that CA has developed over the past 40 years provides a basis for other areas of research that are concerned with social interaction. The relationships that we have identified between talk and non-vocal practices are relevant to and present a challenge to a broad range of gesture studies research from fields such as psycholinguistics, neuropsychology, sociolinguistics, and interactional linguistics. For example, a neuropsychologist may be interested to study what happens in the brain when responding with a relevant gesture, or receiving a relevant gestural response, rather than talk.

Using conversation analytic methods, we have established that non-vocal practices can be used instead of talk and that talk, vocalisations, and non-vocal practices can be used to implement similar actions. It is not necessarily news that gestures and other non-vocal practices can occupy important places within sequences, as non-CA work has addressed this previously. However, the potential sequentiality of non-vocal practices has implications for conversation analytic methodology. That interlocutors respond with

gestures as responses or preliminaries to responses when they may have difficulty speaking, producing the implicated response, or being heard supports the notion that interlocutors are sometimes obligated to respond and to respond in a timely manner. We have also seen that there exist some differences between these modalities, differences that participants on occasion exploit to achieve different interactional effects in the local sequential environment. These observations are important to conversation analytic methodology in the consideration of action formation within the context of sequence organisation. Analysis of sequences based on actions rather than turns reveals that considering responsive actions as talk-specific events is an ideological choice, rather than a participant-oriented-to status. By not clearly distinguishing between actions and turns in sequence organisation, and thereby maintaining that these actions are only implemented through talk, conversation analysts could limit CA's efficacy and its relevance to research that actively engages with non-vocal behaviour. It is important to hold to the principle that analysis should be grounded in how participants orient to practices and resources rather than approaching interaction with preconceived notions of the primacy of talk.

## **Chapter 9**

### **Conclusion: Developing a Comprehensive Conversation Analytic Methodology**

### Abstract

This thesis has addressed the important issue of how conversation analysis, a traditionally talk-focused methodology, can meaningfully incorporate participants' orientations to non-vocal practices. Although there is a growing body of research involving non-vocal practices that occur alongside talk, there is little conversation analytic research that looks at non-vocal practices that occur without accompanying talk. Fundamentally, the difficulty is a researcher orientation to silence as indicative of trouble, whereas interactants might not have the same orientation. By examining silence and non-vocal practices in primarily two settings, television-watching and psychotherapy, that are structured through silence and talk respectively, I have shed light on how silence and stand-alone non-vocal practices operate. The meanings and appropriateness of silence can differ by context, and some contexts can involve a reversal of preference structures with regard to silence. Organisations can also differ for non-vocal and vocal resources. Although talk and non-vocal practices differ with regard to some organisations, sequences of action and timing of action onset appear to be similar between these modalities. By clearly distinguishing between *actions* and *turns* within sequences, the status of non-vocal practices as constituting actions within sequences becomes much more appropriate than relying on talk. Sensitivity to these issues and to participants' orientations is necessary for meaningful analysis of co-present data.

**[1] Introduction**

Despite the presence of conversation analytically informed work on gestures that occur concurrently with talk, little work has been done on non-vocal practices that occur without accompanying talk. On the contrary, classical treatments of conversational organisations, particularly sequence organisation and turn-taking, have been clear that talk is the object of study (e.g., Schegloff, 2007). Silence has been treated as an indicator of interactional trouble or dispreference (e.g., Pomerantz, 1984a; Jefferson, 1988). An initial treatment of silence as problematic is understandable, given the origins of CA being with telephone conversations. However, co-present interactions involve a process of mutual monitoring of not only who is speaking but who is engaged, potentially engaged, and undesirable to have engaged in the encounter. Telephone talk ‘must first be seen as a departure from the norm, else its structure and significance will be lost’ (Goffman, 1964). Work on differences in silence has primarily focused on language differences (e.g., Stivers et al., 2009), but Gardner, Fitzgerald, and Mushin (2009) demonstrate that silence differs contextually and propose that what appear to be cultural differences may in fact be different orientations to context that vary by culture. By treating silence, including that during which other relevant behaviour occurs, as fundamentally problematic, CA becomes unable to cope not only with contextual and cultural differences but also with the reality of co-present interaction in which substantive communication and communication-relevant behaviours do not only occur during talk.

Through a series of studies involving silence and non-vocal practices, I have aimed to address how a more comprehensive approach can benefit conversation analysis. Philips (1985) describes situations as structured through silence or structured through talk, and the different ways in which situations are oriented to by participants may include silence and non-vocal practices. As Stivers and Rossano (2010) argue, treating interactions that are

structured through talk as the point of reference for interactional research is problematic in that most of day-to-day human interaction does not involve long stretches of talk. I have primarily drawn from data involving participants who are engaging in psychotherapy and data involving familiars watching television or films at home. These are two interactional environments that are structured through talk and silence, respectively, and thus radically different orientations to talk might be expected. Such differences in orientation have been demonstrated in this thesis. In this corpus, television-watching involves significant participant silence and stand-alone gesture as well as relaxed sequence organisational considerations, whereas psychotherapy involves significant silence, some stand-alone gesture, and strong sequence organisational features that nonetheless differ from canonical conversations. These differing orientations shed light on how silence and stand-alone gesture operate during co-present interaction and how analytic consideration of context-specific adaptation of interactional structures is necessary to fully understand participant orientations in the local sequential environment.

Jefferson's (1988) suggestion that there is a standard maximum silence of one second in conversation has been the subject of much debate in CA and other fields. Following primarily from Gardner et al.'s (2009) work on co-present speakers of Australian English and Garrwa, I have shown how co-present speakers of American English may exceed Jefferson's standard maximum silence as well. I have then explored a context in which silence, rather than signaling a forthcoming dispreferred response, signals a forthcoming preferred response. In at least some psychotherapeutic contexts, silence is the mark of a preferred response, despite psychotherapeutic and conversation analytic literature that commonly ascribes negativity to silence. When silence occurs following a therapeutic intervention in my data, it is usually treated as an indicator of sincerity, thoughtfulness, and consideration, whereas immediacy is a marked activity (set apart from routine behaviour as

doing something special). That is, silence in this context indicates a preferred response is forthcoming, whereas immediacy on the part of the client is treated as challenging, insincere, or otherwise problematic.

Sometimes researchers attempt to account for unproblematic silence in co-present interactions by applying concepts such as ‘incipient talk’ or ‘continuing states of incipient talk’. Although the fundamental notion behind these concepts is that talk can begin again at any time, there are many very different and often contradictory uses in the literature. One context that is sometimes described as ‘incipient talk’ or a ‘continuing state of incipient talk’ is that of watching television with others, and Couper-Kuhlen (2010) has claimed that during ‘incipient talk’, sequence organisation does not apply. I have examined television- and film-watching from a participant-oriented perspective, showing that although the obligation to respond is relaxed in this context, both sequence organisation and turn-taking are strongly oriented-to by participants.

During television-watching, gestures are quite common. This is easily understandable, as gestures do not interfere with hearing or speaking (cf. Talsma, Doty, Strowd, & Woldorff, 2006). Gestures and other non-vocal practices can also be used to perform sensitive interactional work, such as downplaying a misunderstanding or thanking instead of praising. Consistent with a conversation analytic approach, I have endeavoured to remain loyal to the participants’ orientations throughout the thesis but particularly with regard to how non-vocal practices are used. Because these practices may be used to perform quite nuanced actions, taking an emic perspective (Pike, 1954/1967) is particularly important to understanding the meanings of non-vocal practices. Even gestures that have been catalogued for multiple cultures (e.g., Morris, 1979, 1994) or that are considered universal can be performed very differently (Pike, 1982). Despite differences in how

practices are carried out, their placement within sequences can be examined systematically. By maintaining a distinction between *action* and *turn* during analysis, I have demonstrated how non-vocal practices, whether communicative or task-based, can occupy sequential positions within interactions between humans.

## **[2] A Critical Evaluation**

Certainly in reviewing conversation analytic, silence, and gesture literature, I have omitted areas that are peripheral to the current thesis. I hope that in selecting what to include and what to exclude, I have maintained the clear focus that is intended. In structuring the thesis, I have progressed from silence as a methodological issue in itself – looking at interpretation and transcription, silence as a resource, and its role in defining interactional states – to its intersection with non-vocal practices, including gesture. It is beyond the scope of this thesis to provide an in depth analysis of all possible incarnations of non-vocal behaviour in social interaction. I have instead focused on this intersection in order to address salient issues in CA that revolve around a talk bias that discounts actions that are implemented through non-vocal practices.

The generalisability of particular practices is limited due to the limited dataset from which these studies have drawn. It is not, however, the aim of most conversation analytic research to state that a particular practice is used everywhere by everybody. I have reported a number of practices that people can use and that may be recognised at least in some contexts. My focus has been, though, on the structural resources available to interactants, for example the possibility of using a gestural response to fulfil the obligation to respond but to avoid giving a response of the implicated type. By doing this with a gesture, the withholding of the implicated type is less obvious than it would be if implemented through talk.



**[3] Original Contributions**

I began the empirical chapters of this thesis with a study that considered tolerance for silence in English, drawing on previous research from the international arena. This study specifically supported previous research by Gardner et al. (2009) that demonstrated contextual differences in the presence of silence in Australian English and Garrwa. Differing occurrences of silence can lead to misinterpretation of data, and I explored four possible ways to balance accuracy of meaning with accuracy of the recording in transcription of data with longer than canonical silences.

I then looked at a specific context in which silences are expected more than in other interactions using data from psychotherapy sessions. This study established the regularity of silence following overtly therapeutic questions. Although preferred responses in ordinary conversation regularly occur without significant gaps, prompt responses to overtly therapeutic questions were treated as problematic or contentious. I examined the minority of therapeutic questions that were responded to without gaps and found that immediacy could be accounted for on other structural grounds or as a disaffiliative device. One of these cases was the beginning of a long, problematic exchange, which was then analysed in its entirety. During this analysis, I showed how immediacy can be treated as insincere and how silence contributes to the performance of sincerity. I also demonstrated how a therapeutic question was used to redirect the session to a more aligned state, possibly by virtue of the obligation to respond to interrogatives with the requested information.

Frequent silences have sometimes been cited as indicative of a ‘continuing state of incipient talk’, and Chapter 5 was an empirical study of definitions and usage of this concept, which is sometimes used as an explanation for phenomena that do not fit canonical talk-biased expectations of conversation. I showed that ‘incipient talk’ is not

applied consistently and explored some common themes in its usage and definitions, bringing in possibly related constructs from other interactionally based disciplines. Because of the lack of empirical evidence that people orient to these ‘continuing states of incipient talk’, I suggest that widespread descriptive research of contexts that are proposed to involve ‘incipient talk’ be undertaken. I have attempted this kind of research in Chapter 6, looking at interactions involving people watching television and films together at home. One of the themes identified was that sequence organisation is purported not to apply in a ‘continuing state of incipient talk’, which would account for the relaxed obligation to respond that was identified in interactions during television- and film-watching. I establish this as a problematic approach and examine sequence organisation and turn-taking in the context of watching television and film. I report a relaxation of response relevance and find support for the organisation of actions within sequences and for turn-taking.

I then explored another context in which initiating actions are met with silence – when responsive actions are built entirely from gesture. By examining gestures in responsive positions, I show how interactants can use gestures to do sensitive interactional work. Four main uses emerge: stance displays in second position, using one action type to do another, bracketing intervening talk, and showing that talk is forthcoming. Although gestures to show that a response is forthcoming have been described elsewhere (Dickerson, 2007), I have shown how potentially problematic actions, involving for example disaffiliation or self-praise, can be downplayed. By using a gestural response to fulfil the obligation to respond but to withhold giving a response of the implicated type, the associated disaffiliation is less obvious than it would be if implemented through talk. Likewise, by using a gesture to show stance in a responsive position, the referent is less clear, thus interpretable as a general stance rather than specifically targeted at a person.

Previous research has examined the internal structure of non-vocal practices (e.g., Sacks & Schegloff, 2002; Lerner, Zimmerman, & Kidwell, 2011; Cibulka, 2012), the sequential placement of particular practices (e.g., Seo & Koshik, 2011; Mortensen, 2012; Whitehead, 2011), and the use of non-vocal practices to complete actions that were begun through talk (Olsher, 2004). This thesis is the first to examine the general applicability of sequence organisation for understanding non-vocal practices. In considering non-vocal practices in sequence organisation, I distinguish between communicative non-vocal practices and tasks for the purpose of describing stand-alone practices in terms of their pragmatic functions. I have established that non-vocal responses must occur at the same time as a vocal response could in order to avoid pursuit or increments, as inaction does not project which kind of action is *not* being taken. I have shown that non-vocal practices can perform sequential actions and that in order to maintain a comprehensive approach to interaction, one must distinguish between actions and turns not only in principle but particularly *during analysis* as well.

#### **[4] Implications and Suggestions for Future Research**

In order to accurately analyse interactions on participants' terms, silences cannot a priori be ascribed meaning based on length. Looking at the patterns of silence in preferred responses, initiating actions, etc. is necessary for understanding whether silence is 'silence' in possibly problematic turns or simply in line with the flow of conversation for the participants. By being careful about transcription practices and being clear about the meaningful or mundane nature of silences in terms of how participants regularly speak, analytical problems can be reduced. Participants do not necessarily speak the same across situations. What is a meaningful silence in one context may be typical of another. Silence is performative and can be interpreted in multiple, context-specific ways. These may include cultural considerations, whether the culture is institutionally or socially based. It

can also depend on the sequential environment in which the silence occurs. One environment that differs from ordinary sequences previously identified in conversational data is that of the therapeutic question in psychotherapy. Although many authors have suggested that silence is attributable to the client, engenders resistance, or is even a sign of psychopathology, I have shown that a *lack* of silence may be treated as problematic in situ and that clients may use immediacy to do disaffiliative actions. By adjusting their perceptions of silence and reducing their discomfort with lengthy gaps and pauses (which could stem from being skilled in ordinary conversation), psychotherapists and counsellors could become more attuned to clients and further the therapeutic agenda more effectively. The current research could contribute to training in counselling and psychotherapy communication as well as in other clinical fields. Further research on properties of silence in other potentially sensitive settings and therapeutic approaches could enhance the knowledge that the current research has provided.

Another sequential environment in which silence has been implicated is that of the ‘continuing state of incipient talk’. I have shown that such a state is not empirically grounded. Recommendations include coordinated descriptive research to explore how a large variety of situations that are proposed to constitute ‘incipient talk’ actually are brought about. I have done just this with regard to aspects of sequence organisation during television- and film-watching. Although relevance rules are relaxed during this activity, neither the overall organisation of actions within sequences nor turn-taking are affected. By not using an unsubstantiated construct to explain behaviour, but instead focusing on participant orientation, conversation analysts and other interactional researchers can engage more fully with phenomena that may be found in interactions that are structured through silence.

Finally, the use of non-vocal practices without accompanying talk has been little-researched from a conversation analytic perspective. Because the original organisations that were identified were identified through the study of audio-recordings, there has been limited application to non-vocal practices. I have established that participants indeed treat non-vocal practices as occupying sequential positions and that some properties of turn-taking apply not only to talk but also to non-vocal practices, in particular action onset. However, others (e.g., simultaneous occurrence, or ‘overlap’ as it is termed for talk) are different between the two modalities. The current research has focused specifically on actions within clearly defined sequences, and further research is needed to identify even more nuanced ways in which non-vocal and vocal resources interact. Despite a predominately talk-centric approach to interaction, CA offers an ideal framework for exploring these issues. This thesis has argued how the CA framework can accommodate the study of non-vocal resources that occur both with and without vocal resources. Primarily, this can be accomplished by identifying *actions* in sequences, rather than *turns*, when sequentially analysing data. Sensitivity to contextual meanings of silence, including possibly non-meaningfulness, and what other behaviours occur during silence is necessary to accurately analyse co-present data.

### References

- Alam, S. (2007). *Reframing the majority world*. Retrieved from [http://auroraforum.stanford.edu/files/transcripts/Aurora\\_Forum\\_Essay\\_Shahidal%20Alam\\_Reframing%20the%20Majority%20World.10.01.07.pdf](http://auroraforum.stanford.edu/files/transcripts/Aurora_Forum_Essay_Shahidal%20Alam_Reframing%20the%20Majority%20World.10.01.07.pdf)
- Ansara, Y.G., & Hegarty, P. (2012). Cisgenderism in psychology: pathologising and misgendering children from 1999 to 2008. *Psychology & Sexuality*, 3(2), 137-160.
- Aoki, Hiromi, 2011. Some functions of speaker head nods. In J. Streeck, C. Goodwin, & C. LeBaron (Eds.), *Embodied interaction: Language and body in the material world* (pp. 93-105). Cambridge: Cambridge University Press.
- Arber, A. (2007). "Pain talk" in hospice and palliative care team meetings: An ethnography. *International Journal of Nursing Studies*, 44(6), 916-926.
- Arden, G.B. (1997). Amblyopia. In E. Auerbach (Ed.), *Experimental and clinical amblyopia* (pp. 31-38). Berlin: Springer.
- Austin, G. (1806). *Chironomia; or, a treatise on rhetorical delivery*. London: W. Bulmer, and Co.
- Baker, S. (1955). The theory of silences. *The Journal of General Psychology*, 53, 145-167.
- Baldauf, H. (2002). *Knappes sprechen*. Tübingen: Niemeyer.
- Bales, R.F. (1950). *Interaction process analysis: A method for the study of small groups*. Cambridge, MA: Addison-Wesley.
- Bales, R.F. (1969). *Personality and interpersonal behavior*. Austin, TX: Holt, Rinehart, and Winston.
- Bateson, G. (1935). Culture contact and schismogenesis. *Man*, 35, 178-183.
- Bateson, G. (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology*. Chicago: University Of Chicago Press.

- Bavelas, J.B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction*, 27(3), 201-221.
- Berger, I. (2011). Support and evidence for considering local contingencies in studying and transcribing silence in conversation. *Pragmatics*, 21, 291-306.
- Berger, I. & Rae, J. (2012). Some uses of gestural responsive actions. *Journal of Pragmatics*, 44(13), 1821-1835.
- Berger, I., & Rae, J. (under review/Chapter 8). *Are non-vocally implemented actions subject to the same organisational constraints as talk-implemented actions?*
- Berger, I., & Rae, J. (2010, July 7) *On building responsive actions through non-vocal practices*. Paper presented at the 3rd International Conference on Conversation Analysis, Mannheim, Germany.
- Berger, I. & Viney, R. (2010). *Talking during films: Differing treatments of absent and gestural responses by first speaker*. Paper presented at the 4th Conference of the International Society for Gesture Studies, Frankfurt (Oder), Germany.
- Birdwhistell, R.L. (1952). *Introduction to kinesics: An annotation system for analysts of body motion and gesture*. Louisville, KY: University of Kentucky.
- Birdwhistell, R.L. (1968). Kinesics. In D.L. Sills (Ed.), *International encyclopedia of the social sciences* (Volume 8, p. 379). New York: Macmillan.
- Birdwhistell, R.L. (1970). *Kinesics and context: Essays in body motion communication*. Philadelphia: University of Pennsylvania Press.
- Billig, M. (1999a) Whose terms? Whose ordinariness? Rhetoric and ideology in conversation analysis. *Discourse and Society*, 10, 543-558.
- Billig, M. (1999b). Conversation analysis and the claims of naivety. *Discourse and Society*, 10, 572-576.

- Bloch, S., & Wilkinson, R. (2004). The understandability of AAC: A conversation analysis study of acquired dysarthria. *Augmentative and Alternative Communication*, 20(4), 272-282.
- Blos, P. (1972). Silence: a clinical exploration. *Psychoanalytic Quarterly*, 41, 318-363.
- Boden, D. (1986). Talking with doctors: Conversation analysis in action. *Contemporary Sociology*, 15(5), 715-718.
- Bond, C.F., & DePaulo, B.M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214-234.
- Boomer, D.S. (1965). Hesitation and grammatical encoding. *Language and Speech*, 8(3), 148-158.
- Braithwaite, C.A. (1990). Communicative silence: A cross-cultural study of Basso's hypothesis. In D. Carbaugh (Ed.), *Cultural communication and intercultural contact* (pp. 321-327). Hillsdale, NJ: Lawrence Erlbaum.
- Brooker, B., & Jermyn, D. (2002). *The audience studies reader*. Abingdon, UK: Routledge.
- Brown, P., & Levinson, S. (1978). Universals in language usage: Politeness phenomena. In E. Goody (Ed.), *Questions and politeness* (pp. 56-289). Cambridge: Cambridge University Press.
- Butterworth, B., & Hadar, U., (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, 96, 168-174.
- Calbris, G. (1990). *Semiotics of French gesture*. Bloomington: Indiana University Press.
- Chafe, W. (1980). Some reasons for hesitating. In W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 169-180). The Hague: Mouton.



- Cibulka, P. (2012, March 22). *Home position revisited – Some interactional functions of gesture*. Paper presented at Discourse-Communication-Conversation, Loughborough, UK.
- Clift, R. (2001). Meaning in interaction: The case of 'actually'. *Language*, 77, 245-291.
- Collins, S., & Marková, I. (1995). Complementarity in the construction of a problematic utterance in conversation. In I. Marková, C.F. Graumann, & K. Foppa (Eds.), *Mutualities in dialogue* (pp. 238-263). Cambridge: Cambridge University Press.
- Cook, J.J. (1964). Silence in psychotherapy. *Journal of Counseling Psychology*, 11(1), 42-46.
- Couper-Kuhlen, E. (2010). Commentary on Stivers and Rossano: Mobilizing response. *Research on Language & Social Interaction*, 43, 32-37.
- Coupler-Kuhlen, E., & Ford, C. E. (Eds.). (2004). *Sound patterns in interaction*. Amsterdam: John Benjamins.
- Coupland, J., Robinson, J. D., & Coupland, N. (1994). Frame negotiation in doctor-elderly patient consultations. *Discourse and Society*, 5(1), 89-124.
- Coupland, N., & Coupland, J. (1997). Discourses of the unsayable: Death-implicative talk in geriatric medical consultations. In A. Jaworski (Ed.), *Silence: Interdisciplinary perspectives* (pp. 117-152). Berlin: Mouton de Gruyter.
- Davies, A. (1989). Communicative competence as language use. *Applied Linguistics*, 10(2), 157-170.
- Davies, A. (2007). *Contemplating silence: A review of understandings and clinical handling of patient silence in psychoanalytic psychotherapy* (MHSc dissertation). Auckland University of Technology, Auckland.
- de Fornel, M. (1992). The return gesture: Some remarks on context, inference, and iconic gesture. In P. Auer, & A. Di Luzio (Eds.), *The contextualization of language* (pp. 159-176). Amsterdam: John Benjamins.

- Dickerson, P. (2007). *'Doing thinking': Gaze orientation in post question silence as a means of projecting relevant next action*. (Unpublished manuscript).
- Dickerson, P., Stribling, P., & Rae, J. (2007). Tapping into interaction: How children with autistic spectrum disorders design and place tapping in relation to activities in progress. *Gesture*, 7(3), 271-303.
- Drake and Ford (2011, July 4). *Bodily-Visual Action in Pursuit of Response*. Paper presented at the 12<sup>th</sup> International Pragmatics Conference. Manchester, UK.  
Accessible from  
<http://ipra.ua.ac.be/download.aspx?c=.CONFERENCE12&n=1422&ct=1295&e=10200>
- Drew, P. (1997). "Open" class repair initiators in response to sequential sources of troubles in conversation'. *Journal of Pragmatics*, 28, 69-101.
- Dynel, M. (2011). Revisiting Goffman's postulates on participant statuses in verbal interaction. *Language and Linguistics Compass*, 5, 454-465.
- Eades, D. (2000). I don't think it's an answer to the question: Silencing Aboriginal witnesses in court. *Language in Society*, 29(2), 161-195.
- Eades, D. (2007). Understanding Aboriginal silence in legal contexts. In H. Kotthoff & H. Spencer-Oatey (Eds.), *Handbook of intercultural communication* (pp. 285-301). Berlin: Mouton de Gruyter.
- Efron, D. (1941/1972). *Gesture and environment*. New York: King's Crown Press.
- Ekman, P. & Friesen, W.V. (1969a). Nonverbal leakage and clues to deception. *Psychiatry*, 32, 88-105.
- Ekman, P., & Friesen, W.V. (1969b). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1), 49-98.
- Ekman, P., & Keltner, D. (1997). Universal facial expressions of emotion: An old controversy and new findings. In U. Segerstråle & P. Molnár (Eds.), *Nonverbal*

- communication: Where nature meets culture* (pp. 27-46). Mahwah, NJ: Lawrence Erlbaum.
- Engel, J.J. (1785). *Ideen zu einer mimik*. Berlin: August Mylius.
- Ephratt, M. (2008). The functions of silence. *Journal of Pragmatics*, 40, 1909-1938.
- Erickson, F. (1979). Talking down: Some cultural sources of miscommunication of interracial interviews. In A. Wolfgang (Ed.), *Nonverbal behavior: Applications and cultural implications* (pp. 99-126). New York: Academic Press.
- Feldstein, S., Alberti, L., & BenDubba, M. (1979). Self-attributed personality characteristics and the pacing of conversational interaction. In A.W. Seigman & S. Feldstein (Eds.), *Of speech and time: Temporal speech patterns in interpersonal contexts* (pp. 73-87). Hillsdale, NJ: Lawrence Erlbaum.
- Feyereisen, P. (1987). Gestures and speech, interactions and separations: A reply to McNeill (1985). *Psychological Review*, 94, 493-498.
- Fliess, R. (1949). Silence and verbalization. *International Journal of Psychoanalysis*, 30(1), 21-30.
- Ford, C.E., Thompson, S.A., & Drake, V. (2012). Bodily-visual practices and turn continuation. *Discourse Processes*, 49, 192-212.
- Frankel, Z., Levitt, H.M., Murray, D.M., Greenberg, L.S., & Angus, L. (2006). Assessing silent processes in psychotherapy: An empirically derived categorization system and sampling strategy. *Psychotherapy Research*, 16, 627-638.
- Gale, J., & Sanchez, B. (2005). The meaning and function of silence in psychotherapy with particular reference to a therapeutic community treatment programme. *Psychoanalytic Psychotherapy*, 19(3), 205-220.
- Gardner, R. (1997). The conversational object mm: A weak and variable acknowledging token. *Research on Language and Social Interaction*, 30(2), 131-156.

- Gardner, R. (2002). *When listeners talk: Response tokens and listener stance*. Philadelphia: John Benjamins.
- Gardner, R., Fitzgerald, R., & Mushin, I. (2009). The underlying orderliness in turn-taking: Examples from Australian talk. *Australian Journal of Communication*, 36(3), 65-90.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Garfinkel, H., & Rawls, A.W. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Lanham, MD: Rowman & Littlefield.
- Gerhardt, C. (2006). Moving closer to the audience: Watching football on television. *Revista Alicantina de Estudios Ingleses*, 19, 125-148.
- Gerhardt, C. (2009). Multimodal and intertextual humor in the media reception situation: The case of watching football on TV. In N. R. Norrick & D. Chiaro (Eds.), *Humor in interaction* (pp. 79-100). Amsterdam: John Benjamins.
- Giles, H., and J.M. Wiemann (1987) Language, social comparison and power. In C.R. Berger and S.H. Chaffee (eds.), *The handbook of communication science* (pp. 350-384). Newbury Park, CA: Sage.
- Goffman, E. (1955). On face-work: An analysis of ritual elements of social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes*, 18(3), 213-231.
- Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gatherings*. New York: The Free Press.
- Goffman, E. (1964). The neglected situation. *American Anthropologist*, 66(6), 133-136.
- Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. Garden City, NJ: Anchor.
- Goffman, E. (1971). *Relations in public: Microstudies of the public order*. New York: Basic Books.
- Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97-121). New York: Irvington.
- Goodwin, C. (1980). Restarts, pauses, and the achievement of mutual gaze at turn-beginning. *Sociological Inquiry*, 50(3-4), 272-302.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Goodwin, C. (2003). Conversational frameworks for the accomplishment of meaning in aphasia. In C. Goodwin (Ed.), *Conversation and brain damage* (pp. 90-116). Oxford: Oxford University Press.
- Goodwin, C. (2004). A competent speaker who can't speak: The social life of aphasia. *Journal of Linguistic Anthropology*, 14(2), 151-170.
- Goodwin, C. (2006). Human sociality as mutual orientation in a rich interactive environment: Multimodal utterances and pointing in aphasia. In N. Enfield and S. C. Levinson (Eds.), *Roots of human sociality* (pp. 96-125). London: Berg Press.
- Goodwin, C. (2007). Participation, stance and affect in the organization of activities. *Discourse Society*, 18, 53-73.
- Goodwin, C. (2011). Constructing meaning through prosody in aphasia. In D. Barth-Weingarter, E. Reber, & M. Selting (Eds.), *Prosody in interaction* (pp. 373-394). Amsterdam: John Benjamins.
- Goodwin, C., & Goodwin, M.H. (1986). Gesture and coparticipation in the activity of searching for a word. *Semiotica*, 62(1-2), 51-75.
- Goodwin, C., & Goodwin, M.H. (1987). Concurrent operations on talk: Notes on the interactive organization of assessments. *IPrA Papers in Pragmatics*, 1(1), 1-54.

- Goodwin, C., & Goodwin, M.H. (1992). Context, activity and participation. In P. Auer & A. Di Luzio (Eds.), *The contextualization of language* (pp. 77-99). Amsterdam: John Benjamins.
- Goodwin, M.H. (1980). Processes of mutual monitoring implicated in the production of description sequences. *Sociological Inquiry*, 50(3-4), 303-17.
- Goodwin, M.H. (1983). Searching for a word as an interactive activity. J.N. Deely & M.D. Lenhart (Eds.), *Semiotics* (pp. 129-138). New York: Plenum.
- Goodwin, M.H. (1985). The serious side of jump rope: Conversational practices and social organization in the frame of play. *Journal of American Folklore*, 98, 315-330.
- Goodwin, M.H. (1998). Games of stance: Conflict and footing in hopscotch. In S. Hoyle and C.T. Adger (Eds.), *Kids' talk: Strategic language use in later childhood* (pp. 23-46). New York: Oxford University Press.
- Goodwin, M.H. (2006). *The hidden life of girls: Games of stance, status, and exclusion*. Oxford: Blackwell.
- Goodwin, M.H., & Goodwin, C. (2001). Emotion within situated activity. In A. Duranti (Ed.), *Linguistic anthropology: A reader* (pp. 239-257). Malden, MA: Blackwell.
- Goodwin, M.H., Goodwin, C., & Yaeger-Dror, M. (2002). Multi-modality in girls' game disputes. *Journal of Pragmatics*, 34, 1621-1649.
- Gourley, L.E. (1973). Capital B for Black. *American Psychologist*, 30(2), 181.
- Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of French and Swedish*. Lund, Sweden: Lund University Press.
- Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text & Talk*, 26, 281-328.
- Haim, R.J. (1990). The timing of interventions: A countertransference dilemma, when to talk and when not to talk. *Modern Psychoanalysis*, 15, 79-87.
- Hall, E. (1969). *The hidden dimension*. Garden City, NJ: Double Day.

- Hawkins, P.R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, 14(3), 277-288.
- Hayashi, M. (2010). An overview of the question-response system in Japanese. *Journal of Pragmatics*, 42, 2685-2702.
- Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge: Cambridge University Press.
- Heath, C. (1992). Gesture's discreet tasks: Multiple relevancies in visual conduct and in the contextualisation of language. In P. Auer & A. Di Luzio (Eds.), *The contextualization of language* (pp. 101-127). Amsterdam: John Benjamins.
- Heath, C.C., & Luff, P. (1993). Disembodied conduct: Communication through video in a multi-media office environment. In G. Button (Ed.), *Technology in working order: Studies of work, interaction, and technology* (pp. 35-54). London: Routledge.
- Helm, D.T. (1982). Talk's form: Comments on Goffman's "Forms of Talk". *Human Studies*, 5(2), 147-157.
- Heritage, J. (1984a). *Garfinkel and ethnomethodology*. Cambridge: Polity.
- Heritage, J. (1984b). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299-345). Cambridge, UK: Cambridge University Press.
- Heritage, J. (1998). Oh-prefaced responses to inquiry. *Language in Society*, 27, 291-334.
- Heritage, J. (2005). Revisiting authority in physician-patient interaction. In J. Duchan & D. Kovarsky (Eds.) *Diagnosis as cultural practice* (pp. 83-102). New York: Mouton De Gruyter.
- Heritage, J., & Atkinson, J.M. (1984). Introduction. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 1-15). Cambridge: Cambridge University Press.

- Heritage, J., & Maynard, D.W. (2006). Introduction: Analyzing interaction between doctors and patients in primary care encounters. In J. Heritage & D. W. Maynard (Eds.), *Communication in medical care: Interaction between primary care physicians and patients* (pp. 1-21). Cambridge: Cambridge University Press.
- Heritage, J., & Maynard, D.W. (Eds.) (2006). *Communication in medical care: Interaction between primary care physicians and patients*. Cambridge: Cambridge University Press.
- Heritage, J. & Robinson, J.D. (2006). Accounting for the visit: Giving reasons for seeking medical care. In J. Heritage & D. Maynard (Eds.), *Communication in medical care: Interaction between primary care physicians and patients* (pp. 48-85). Cambridge: Cambridge University Press.
- Heritage, J., Robinson, J.D., Elliott, M., Beckett, M., & Wilkes, M. (2007). Reducing patients' unmet concerns in primary care: The difference one word can make. *Journal of General Internal Medicine*, 22, 1429-1433.
- Heritage, J., & Roth, A. (1995). Grammar and institution: Questions and questioning in the broadcast news interview. *Research on Language and Social Interaction*, 28(1), 1-60.
- Hill, C.E. (1978). Development of a counselor verbal response category system. *Journal of Counseling Psychology*, 25, 461-468.
- Hill, C.E., Thompson, B.J., & Ladany, N. (2003) Therapist use of silence in therapy: A survey. *Journal of Clinical Psychology*, 59, 513-524.
- Holt, E. (2010). The last laugh: Shared laughter and topic termination. *Journal of Pragmatics*, 42(6), 1513-1525.
- Holt, E., & Drew, P. (1995). Idiomatic expressions and their role in the organization of topic transition in conversation. In M. Everaert, E. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 117-131). Hillsdale, NJ: Lawrence Erlbaum.



- Homans, G.C. (1958). Social behavior as exchange. *American Journal of Sociology*, 63, 597-606.
- Ide, S. (1989). Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. *Multilingua* 8(2/3), 223-248.
- Ishii, S., & Bruneau, T. (1988). Silence and silences in cross-cultural perspective: Japan and the United States. In L. Samovar & P. Porter (Eds.), *Intercultural communication* (pp. 246-251). Belmont, CA: Wadsworth.
- Iverson, J.M., & Goldin-Meadow, S. (2001). The resilience of gesture in talk: Gesture in blind speakers and listeners. *Developmental Science*, 4, 416-422.
- Jefferson, G. (1972). Side sequences. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 294-333). New York: Free Press.
- Jefferson, G. (1974). Error correction as an interactional resource. *Language in Society*, 3(2), 181-199.
- Jefferson, G. (1975/2004). A sketch of some orderly aspects of overlap in natural conversation. In G. H. Lerner (Ed.) *Conversation analysis: Studies from the first generation* (pp.43-59). Philadelphia: John Benjamins.
- Jefferson, G. (1978). What's in a 'nyem'? *Sociology*, 12, 135-139.
- Jefferson, G. (1983). Issues in the transcription of naturally-occurring talk: Caricature versus capturing pronunciation particulars. *Tilburg Papers in Language and Literature*, 34, 1-12.
- Jefferson, G. (1984a). On stepwise transition from talk about a trouble to inappropriately next-positioned matters. In J. M. Atkinson and J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge: Cambridge University Press.

- Jefferson, G. (1984b). Notes on some orderlinesses of overlap onset. In V. D'Urso and P. Leonardi (Eds.), *Discourse analysis and natural rhetoric* (pp. 11-38). Padua, Italy: Cleup Editore.
- Jefferson, G. (1986). Notes on 'latency' in overlap onset. *Human Studies*, 9, 153-183.
- Jefferson, G. (1988). Notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. Roger and P. Bull (Eds.), *Conversation: An interdisciplinary perspective* (pp. 167-196). Clevedon, UK: Multilingual Matters.
- Johannesen, R.L. (1974). The functions of silence: A plea for communication research. *Western Speech*, 38(1), 25-35.
- Jones, R. & Thornborrow, J. (2004). Floors, talk and the organization of classroom activities. *Language in Society* 33, 399-423.
- Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Kendon, A. (1982). Some relationships between body motion and speech. In A. Seigman and B. Pope (Eds.), *Studies in dyadic communication* (pp. 177-216). Elmsford, NY: Pergamon Press.
- Kendon, A. (1985). Some uses of gesture. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 215-234). Norwood, NJ: Ablex.
- Kendon, A. (1988a). Goffman's approach to face-to-face interaction. In P. Drew (Ed.), *Erving Goffman: Exploring the interaction order* (pp. 14-40). Chicago: Northeastern University Press.
- Kendon, A. (1988b). How gestures can become like words. In F. Poyatos (Ed.), *Crosscultural perspectives in nonverbal communication* (pp. 131-141). Toronto: Hogrefe.

- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
- Kendon, A. (1992). Some recent work from Italy on quotable gestures (emblems). *Journal of Linguistic Anthropology*, 2(1), 92-108.
- Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, 27(3), 175-200.
- Kendon, A. (1997). Gesture. *Annual Review of Anthropology*, 26, 109-128.
- Kendon, A. (2002). Some uses of the head shake. *Gesture*, 2(2), 147-182.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A., & Sigman, S.J. (1996). Ray L. Birdwhistell (1918-1994). *Semiotica*, 112(3/4), 231-261.
- Kern, F. (2010) Speaking dramatically: The prosody of live radio commentary of football matches. In D. Barth-Weingarten, E. Reber, & M. Selting (Eds.), *Prosody in interaction* (pp. 217-238). Amsterdam: John Benjamins.
- Kidwell, M. (2005). Gaze as social control: How very young children differentiate “the look” from a “mere look” by their adult caregivers. *Research on Language and Social Interaction*, 38(4), 417-449.
- Kidwell, M. (2010, July 8). *Recognizability in the interactions of very young children: Showing vs. giving*. Paper presented at the 3rd International Conference on Conversation Analysis. Mannheim, Germany.
- Kidwell, M., & Zimmerman, D. (2007). Joint attention as action. *Journal of Pragmatics* 39, 592-611.
- Kim, M.-S. (2002). *Non-western perspectives on human communication: Implications for theory and practice*. Thousand Oaks, CA: Sage.

- Kincaid, D.L. (1979). *The convergence model of communication*. East-West Communication Institute Paper 18. Honolulu, HI: East-West Center.
- Kitzinger, C. (2008). Conversation analysis: Technical matters for gender research. In K. Harrington, L. Litosseliti, H. Sauntson, & J. Sunderland (Eds.), *Gender and language: Theoretical and methodological approaches* (pp. 119-138). Basingstoke: Palgrave Macmillan.
- Kitzinger, C. (2011). Working with childbirth helplines: The contributions and limitations of conversation analysis. In C. Antaki (Ed.), *Applied conversation analysis: Intervention and change in institutional talk*. Basingstoke: Palgrave.
- Kitzinger, C. & Kitzinger, S. (2007). Birth trauma: Talking with women and the value of conversation analysis. *British Journal of Midwifery*, 15(5), 256-264.
- Kitzinger, C., Shaw, R., & Toerien, M. (2012). Referring to persons without using a full-form reference: Locally initial indexicals in action. *Research on Language and Social Interaction*, 45, 116-136.
- Kollock, P, Blumstein, P., & Schwartz, P. (1985). Sex and power in interaction: Conversational privileges and duties. *American Sociological Review*, 50(1), 34-46.
- Korkiakangas, T. (2011). *Eye-gaze in Multimodal Interactions involving children with Autism Spectrum Disorders* (Doctoral thesis). London: University of Roehampton.
- Korkiakangas, T. (2012). *Dementia, interaction and identity*. Paper presented at Among Others: Different Approaches to Organising Social Action, University of Roehampton, London.
- Koschmann, T., & LeBaron, C. (2002). Learner articulation as interactional achievement: Studying the conversation of gesture. *Cognition & Instruction*, 20(2), 249-282.
- Land, V., and Kitzinger, C. (2005). Speaking as a lesbian: Correcting the heterosexist presumption. *Research on Language and Social Interaction*, 38, 371-416.

- Lehtinen, E. (2007). Merging doctor and client knowledge: On doctors' ways of dealing with clients' potentially discrepant information in genetic counseling. *Journal of Pragmatics*, 39(2), 389-427.
- Lehtonen, J., & Sajavaara, K. (1985). The silent Finn. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp.193-201). Norwood, NJ: Ablex.
- Lerner, G.H. (1996). On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In E. Ochs, E.A. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 238-276). Cambridge: Cambridge University Press.
- Lerner, G.H. (2002). *Practice does not make perfect: Intervening actions in the selection of next speaker*. Plenary address at Conference on Language, Interaction and Culture, University of California at Los Angeles.
- Lerner, G.H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society*, 32(2). 177-201.
- Lerner, G.H., & Raymond, G. (2008). *Body trouble: Some sources of interactional trouble and their embodied solution* (Unpublished manuscript).
- Lerner, G.H., Zimmerman, D.H., & Kidwell, M. (2011). Formal structures of practical tasks: A resource for action in the social life of very young children. In J. Streeck, C. Goodwin, & C. LeBaron, *Embodied interaction: Language and body in the material world* (44-58). Cambridge: Cambridge University Press.
- Leudar, I., Sharrock, W., Truckle, S., Colombino, T., Hayes, J., & Booth, K. (2008). Conversation of emotions: On turning play into psychoanalytic psychotherapy. In A. Peräkylä, C. Antaki, S. Vehviläinen, and I. Leudar (Eds.), *Conversation analysis and psychotherapy* (pp. 152-172). Cambridge: Cambridge University Press.
- Levinson, S.C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

- Levitt, H.M. (1998). Silence in psychotherapy: The meaning and function of pauses (Doctoral dissertation). York University, Toronto.
- Levitt, H.M. (2002). Clients' experiences of obstructive silence: Integrating conscious reports and analytic theories. *Journal of Contemporary Psychotherapy*, 31(4), 221-244.
- Lindström, A.K.B. (1997). Designing social actions: Grammar, prosody, and interaction in Swedish conversation (PhD Dissertation). Department of Sociology, University of California, Los Angeles.
- Local, J., & Walker, G. (2004). Abrupt-joins as a resource for the production of multi-unit, multi-action turns. *Journal of Pragmatics* 36, 1375-1403.
- Łojek, E. (2009). Imaging communication in the brain. In K. Bryan (Ed.), *Communication in healthcare* (pp. 1-66). Bern, Switzerland: Peter Lang.
- Macbeth, D. (2004). The relevance of repair for classroom correction. *Language in Society*, 33, 703-736.
- Madill, A., Widdicombe, S., & Barkham, M. (2001). The potential of conversation analysis for psychotherapy research. *The Counseling Psychologist*, 29(3), 413-434.
- Mahl, G.F. (1956). Disturbances and silences in the patient's speech in psychotherapy. *The Journal of Abnormal and Social Psychology*, 53(1), 1-15.
- Marler, P., & Evans, C.S. (1997). Communication signals of animals: Contributions of emotion and reference. In U. Segerstråle & P. Molnár (Eds.), *Nonverbal communication: Where nature meets culture* (pp. 151-170). Mahwah, NJ: Lawrence Erlbaum.
- Matthews, D. (2010). *My computer is so slow: The use of laughables and laughter when dealing with computer problems in call-centre interaction*. Poster presented at the 10th International Conference on Conversation Analysis, Universität Mannheim: Mannheim, Germany.

- Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic Games. *Journal of Personality and Social Psychology*, 91(3), 568-581.
- Matsumoto, Y. (1988). Reexamination of the universality of face: Politeness phenomena in Japanese. *Journal of Pragmatics*, 12, 403-426.
- Maynard, D.W., & Marlaire, C.L. (1992). Good reasons for bad testing performance: The interactional substrate of educational exams. *Qualitative Sociology*, 15(2), 177-202.
- McDermott, R.P., & Tylbor, H. (1983). On the necessity of collusion in conversation. *Text*, 3(3), 277-297.
- McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, (92)3, 350-371.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- McWhorter, J. (2004, September 18) Call me Black. That's with a capital B. Not African-American. *Houston Chronicle*. Retrieved from <http://www.chron.com/dispatch/story.mpl/editorial/outlook/2802824.html>
- Mehrabian, A. (1972). *Nonverbal communication*. New Brunswick, NJ: Aldine.
- Mehrabian, A., & Ferris, S.R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31, 248-252.
- Millward, L. J. (2005). *Understanding occupational and organizational psychology*. London, Sage.
- Mondada, L. (2006). La compétence comme dimension située et contingente, localement évaluée par les participants. *Bulletin VALS-ASLA (Association suisse de linguistique appliquée)*, 84, 83-119.

- Mondada, L. (2009). Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics*, 41, 1977–1997.
- Mondada, L., & Oloff, F. (2010). Gestures in overlap: The situated establishment of speakership. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture*. Amsterdam: John Benjamins.
- Morris, D. (1994). *Bodytalk: A world guide to gestures*. London: Jonathan Cape.
- Mortensen, K. (2012). Visual initiations of repair – Some preliminary observations. In K. Ikeda & A. Brandt (Eds.), *Challenges and new directions in the micro-analysis of social interaction* (pp. 45-50). Osaka: Division of International Affairs, Kansai University.
- Mousavi, S.Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, 87, 319-334.
- Murphy, M. (1970). Silence, the word, and Indian rhetoric. *College Composition and Communication*, 21, 359-360.
- Mushin, I. & Gardner, R. (2009). Silence is talk: Conversational silence in Australian Aboriginal talk-in-interaction. *Journal of Pragmatics*, 41, 2033-2052.
- Nespoulous, J.-L., & Lecours, A. R., (1986). Gestures: Nature and function. In: J. Nespoulous, P. Perron, and A.R. Lecours (Eds.) *The biological foundations of gestures: Motor and semiotic aspects* (pp. 49-62). Hillsdale, NJ: Lawrence Erlbaum.
- Nevile, M. (2004). *Beyond the black box: Talk-in-interaction in the airline cockpit*. Aldershot, UK: Ashgate.
- Nevile, M. (2006). *Communication in context: A conversation analysis tool for examining recorded voice data in investigations of aviation occurrences*. Canberra: Australian Transport Safety Bureau.



- Nevile, M. (2007). Talking without overlap in the airline cockpit: Precision timing at work. *Text & Talk*, 27(2), 225-249.
- Nofsinger, R.E. (1999). *Everyday conversation*. Long Grove, IL: Waveland.
- Nwoye, G.O. (1985). Eloquent silence among the Igbo of Nigeria. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 185-191), Norwood, NJ: Ablex.
- Ochs, E., Gonzales, P., & Jacoby, S. (1996). "When I come down, I'm in the domain state": Grammar and graphic representation in the interpretive activity of physicists. In E. Ochs, E. A. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 1-51). Cambridge: Cambridge University Press.
- Olsher, D. (2004) Talk and gesture: The embodied completion of sequential actions in spoken interaction. In R. Gardner and J. Wagner (Eds.), *Second language conversations* (pp. 221-245). New York: Continuum.
- Pallotti, G. (2001). External appropriations as a strategy for participating in intercultural multi-party conversations. In S. Günthner & F. Orletti (Eds.), *Culture in communication: Analyses of intercultural situations* (pp. 295-334). Amsterdam: John Benjamins.
- Peräkylä, A., Antaki, C., Vehviläinen, S. and Leudar, I. (Eds.) (2008). *Conversation analysis and psychotherapy*. Cambridge: Cambridge University Press.
- Philips, S.U. (1976). Some sources of cultural variability in the regulation of talk. *Language in Society*, 5, 81-95.
- Philips, S.U. (1985). Interaction structured through talk and interaction structured through 'silence'. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 205-213). Norwood, NJ: Ablex.
- Pilnick, A. (2002). There are no rights and wrongs in these situations: Identifying interactional difficulties in genetic counseling. *Sociology of Health and Illness*, 25(1), 66-88.

- Pike, K. L. (1954/1967). *Language in relation to a unified theory of human behavior* (2<sup>nd</sup> Edition). The Hague: Moulton & Co.
- Pike, K.L. (1982). *Linguistic concepts: An introduction to tagmemics*. Lincoln, NE: University of Nebraska Press.
- Pomerantz, A. (1980). Telling my side: 'Limited access' as a 'fishing' device. *Sociological Inquiry*, 50(3/4), 186-198.
- Pomerantz, A. (1984a). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson and J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 57-101). Cambridge: Cambridge University Press.
- Pomerantz, A. (1984b). Pursuing a response. In J. M. Atkinson and J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 152-164). Cambridge: Cambridge University Press.
- Potter, J., & Hepburn, A. (2010). Putting aspiration into words: 'Laugh particles', managing descriptive trouble and modulating action. *Journal of Pragmatics*, 42, 1543-1555.
- Preuschoft, S., & van Hooff, J.A.R.A.M. (1997). The social function of "smile" and "laughter": Variations across primate species and societies. In U. Segerstråle & P. Molnár (Eds.), *Nonverbal communication: Where nature meets culture* (pp. 171-190). Mahwah, NJ: Lawrence Erlbaum.
- Prince, R. (1997). Is talk cheap or silence golden: The negotiation of activity in psychotherapy. *Psychoanalytic Inquiry*, 17(4), 549-558.
- Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. Thousand Oaks, CA: Sage.
- Psathas, G., & Anderson, T. (1990). The 'practices' of transcription in conversation analysis. *Semiotica*, 78, 75-99.

- Rae, J. (2008). Lexical substitution as a therapeutic resource. In A. Peräkylä, C. Antaki, S. Vehviläinen, and I. Leudar (Eds.), *Conversation analysis and psychotherapy* (pp. 62-79). Cambridge: Cambridge University Press.
- Rae, J., & Guimarães, E. (2010). *Passing stuff: How do humans accomplish manual object transfers?*. Paper presented at the 10th International Conference on Conversation Analysis, Universität Mannheim: Mannheim, Germany.
- Rae, J., & Korkiakangas, T. (2011). *Smiling as an interactional resource: A case study of a boy with autism*. Paper presented at the 10th Conference of the International Institute for Ethnomethodology and Conversation Analysis, Université de Fribourg: Fribourg.
- Raymond, G. (2000). The voice of authority: The local accomplishment of authoritative discourse in live news broadcasts. *Discourse Studies*, 2(3), 354-379.
- Raymond, G. (2003). Grammar and social organization: Yes/no interrogatives and the structure of responding. *American Sociological Review*, 68, 939-967.
- Raymond, G. (2010). Prosodic variation in responses: The case of type-conforming responses to yes/no-interrogatives. In D. Barth-Weingarten, E. Reber, & M. Selting, *Prosody in interaction* (pp. 109-130). Amsterdam: John Benjamins.
- Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience*, 27, 169-192.
- Robinson, J. (1998). Getting down to business: Talk, gaze, and body orientation during openings of doctor-patient consultations. *Human Communication Research*, 25(1), 97-123.
- Robinson, J.D., Silk, K. J., Parrott, R. J., Steiner, C., Morris, S. M., & Honeycutt, C. (2004). Healthcare providers' sun-protection promotion and at-risk clients' skin-cancer-prevention outcomes. *Preventive Medicine*, 38, 251-257.
- Rosenblatt, P.C., & Li, X. (2010). Hazards to family relationships from cell phone usage while driving. *Family Science Review*, 15(2), 84-93.

- Rossano, F. (2010). Questioning and responding in Italian. *Journal of Pragmatics*, 42, 2756-2771.
- Rossano, F. (2011). *Communicating without words: The design and recognition of first actions in great apes*. Paper presented at the 12th International Pragmatics Conference. Manchester, UK.
- Ruusuvuori J., & Peräkylä, A. (2009). Facial and verbal expressions in assessing stories and topics. *Research on Language and Social Interaction*, 42(4), 377-394.
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J.R.E. Lee (Eds.), *Talk and social organisation*. Clevedon, UK: Multilingual Matters.
- Sacks, H. (1964/1989/1992). Lecture one: Rules of conversational sequence. *Human Studies*, 12, 333-350. [Also published in *Lectures on conversation, Volumes I and II* (pp. 3-11). Oxford: Blackwell.]
- Sacks, H. (1968/1995). Topic [Lecture 17 April 1968]. Later published in H. Sacks, *Lectures on conversation* (pp. 762-763). Malden, MA: Blackwell.
- Sacks, H. (1992). *Lectures on conversation, Volumes I and II*. Oxford: Blackwell.
- Sacks, H. (1995). *Lectures on conversation*. Oxford: Blackwell.
- Sacks, H., & Schegloff, E.A. (2002). Home position. *Gesture*, 2(2), 133-146.
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696-735.
- Samarin, W. (1965). The language of silence. *Practical Anthropology*, 12, 115-119.
- Sapir, E. (1925). Sound patterns in language. *Language*, 1, 37-51.
- Sapir, E. (1949). Speech as a personality strait. In D. Mandelbaum (Ed.), *Selected writings of Edward Sapir in language, culture, and personality* (pp. 533-543). Berkley: University of California Press.

- Saville-Troike, M. (1985). The place of silence in an integrated theory of communication. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 3-20). Norwood, NJ: Ablex.
- Scheflen, A.E. (1963). Communication and regulation in psychotherapy. *Psychiatry*, 26, 126-136.
- Scheflen, A.E. (1973). *Communicational structure: Analysis of a psychotherapy transaction*. Bloomington: Indiana University Press.
- Scheflen, A.E. (1974). *How behavior means*. New York: Jason Aronson.
- Schegloff, E.A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70, 1075-1095.
- Schegloff, E.A. (1980). Preliminaries to preliminaries: "Can I ask you a question?". *Sociological Inquiry*, 50(3-4), 104-152.
- Schegloff, E.A., (1984). On some gestures' relation to talk. In J. M Atkinson & J. Heritage (Eds.) *Structures of social action: Studies in conversation analysis* (pp. 266-296). Cambridge: Cambridge University Press.
- Schegloff, E.A. (1987). Analyzing single episodes of interaction: An exercise in conversation analysis. *Social Psychology Quarterly*, 50(2), 101-114.
- Schegloff, E.A. (1988). Goffman and the analysis of conversation. In P. Drew & T. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 9-135). Cambridge: Polity.
- Schegloff, E.A. (1990). On the organization of sequences as a source of 'coherence' in talk-in-interaction. In B. Dorval (Ed.), *Conversational organization and its development* (pp. 51-77). Norwood, NJ: Ablex.
- Schegloff, E.A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 97, 1295-1345.

- Schegloff, E.A. (1995). Discourse as an interactional achievement III: The omnirelevance of action. *Research on Language and Social Interaction*, 28(2), 185-211.
- Schegloff, E.A. (1996a). Some practices for referring to persons in talk-in-interaction: A partial sketch of a systematics. In B. Fox (Ed.), *Studies in anaphora* (pp. 437-485). Amsterdam: John Benjamins.
- Schegloff, E.A. (1996b). Issues of relevance for discourse analysis: Contingency in action, interaction and co-participant context. In E.H. Hovy & D. Scott (Eds.), *Computational and conversational discourse: Burning issues – An interdisciplinary account* (pp. 3-38). Heidelberg: Springer.
- Schegloff, E.A. (1996c). Turn organization: One intersection of grammar and interaction. In E. Ochs, E.A. Schegloff and S. Thompson (Eds.), *Interaction and grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E.A. (1998). Body torque. *Social Research*, 65, 535-596.
- Schegloff, E.A. (2004) Answering the phone. In G. H. Lerner (Ed.), *Conversation analysis* (pp. 63-107). Amsterdam: Benjamins.
- Schegloff, E.A. (2005). On complainability. *Social Problems*, 52, 449-476.
- Schegloff, E.A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Schegloff, E.A. (2009). Beginning to respond: Well-prefaced responses to Wh-questions. *Research on Language and Social Interaction*, 42(2), 91-115.
- Schegloff, E.A. (2010). Commentary on Stivers & Rossano: "Mobilizing response". *Research on Language and Social Interaction*, 43, 38-48.
- Schegloff, E.A., Jefferson, G., & Sacks, H., (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361-382.

- Schegloff, E.A., Ochs, E., & Thompson, S. (1996). Introduction. In E. Ochs, E.A. Schegloff, & S. Thompson (Eds.) *Interaction and grammar* (pp. 1-51). Cambridge: Cambridge University Press.
- Schegloff, E.A. & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289-327.
- Scherer, K.R., Feldstein, S., Bond, R.N., & Rosenthal, R. (1985). Vocal cues to deception: A comparative channel approach. *Journal of Psycholinguistic Research*, 14(4), 409-425.
- Scollon, R. (1985). The machine stops. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 21-30). Norwood, NJ: Ablex.
- Scott, V., & Lester, S. (1998). Listening to silence. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 19(3), 105-108.
- Seo, M.-S., & Koshik, I. (2010). A conversation analytic study of gestures that engender repair in ESL conversational tutoring. *Journal of Pragmatics*, 42, 2219-2239.
- Silverman, D. (1996). *Discourses of counselling: HIV counselling as social interaction*. London: Sage.
- Silverman, D., & Peräkylä, A. (1990). AIDS counselling: The interactional organisation of talk about 'delicate' issues. *Sociology of Health & Illness*, 12(3), 293-318.
- Speer, S.A. & Parsons, C. (2006). Gatekeeping gender: Some features of the use of hypothetical questions in the psychiatric assessment of transsexual patients. *Discourse & Society*, 17 (6), 785-812.
- Speer, S.A. & Parsons, C. (2007). Suppose you couldn't go any further with treatment, what would you do? Hypothetical questions in interactions between psychiatrists and transsexual patients. In A. Hepburn & S. Wiggins (Eds.) *Discursive research in practice: New approaches to psychology and everyday interaction*. Cambridge: Cambridge University Press.

- Stifelman, L.J. (1994, September). *The cocktail party effect in auditory interfaces: A study of simultaneous presentation*. Boston, MA: MIT Media Laboratory Technical Report.
- [Accessed from [http://www.media.mit.edu/speech/papers/1994/stifelman\\_TR94\\_cocktail\\_party\\_effect.pdf](http://www.media.mit.edu/speech/papers/1994/stifelman_TR94_cocktail_party_effect.pdf)]
- Stivers, T. (2005a). Parent resistance to physicians' treatment recommendations: Once resource for initiating a negotiation of the treatment decision. *Health Communication*, 18(1), 41-74.
- Stivers, T. (2005b). Non-antibiotic treatment recommendations: Delivery formats and implications for parent resistance. *Social Science and Medicine*, 60(5), 949-964.
- Stivers, T. (2007). *Prescribing under pressure: Parent-physician conversations and antibiotics*. Oxford: Oxford University Press.
- Stivers, T. (2008). Stance, alignment and affiliation during story telling: When nodding is a token of preliminary affiliation. *Research on Language in Social Interaction*, 41, 29-55.
- Stivers, T., Enfield, N. J., & Levinson, S. C. (Eds.) (2010). Question-response sequences in conversation across ten languages [Special Issue]. *Journal of Pragmatics*, 42(10).
- Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J.P., Yoon, K.-E., Levinson, S.C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106, 10587-10592.
- Stivers, T. & Majid, A. (2007). Questioning children: Interactional evidence of implicit racial bias in medical interviews. *Social Psychology Quarterly*, 70, 424-441.
- Stivers, T., Mangione-Smith, R., Elliott, M. N., McDonald, L., & Heritage, J. (2003). What leads physicians to believe that parents expect antibiotics?: A study of parent



- communication behaviors and physicians' perceptions. *The Journal of Family Practice*, 52(2), 140-147.
- Stivers, T., & Robinson, J.D. (2006). A preference for progressivity in interaction. *Language in Society*, 35(3), 367-392.
- Stivers, T., & Rossano, F. (2010a). A scalar view of response relevance. *Research on Language and Social Interaction*, 43, 49-56.
- Stivers, T., & Rossano, F., (2010b). Mobilizing response. *Research on Language & Social Interaction*, 43, 3-31.
- Stokoe, E. (2009). "I've got a girlfriend": Police officers doing 'self-disclosure' in their interrogations of suspects. *Narrative Inquiry*, 19(1), 154-182.
- Streeck, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs*, 60, 275-299.
- Streeck, J. (2002) Grammars, words, and embodied meanings: On the uses and evolution of so and like. *Journal of Communication*, 52, 581-596.
- Streeck, J. (2009a). Forward-gesturing. *Discourse Processes*, 46, 161- 179.
- Streeck, J. (2009b). *Gesturecraft: The manufacture of meaning*. Amsterdam: John Benjamins.
- Streeck, J., & Hartge, U. (1992). Previews: Gestures at the transition place. In P. Auer & A. Di Luzio (Eds.), *The contextualization of language* (pp. 135-157). Amsterdam: John Benjamins.
- Street, R., & Giles, H. (1982). Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. Roloff & C.R. Berger (Eds.), *Social cognition and communication* (pp. 193-226). Beverly Hills: Sage.
- Stribling, P., & Rae, J. (2010). Interactional analysis of scaffolding in a mathematical task. In H. Gardner and M. Forrester (Eds.), *Analysing interactions in childhood: Insights from conversation analysis* (pp. 185-208). Oxford: Wiley Blackwell.

- Stribling, P., Rae, J., & Dickerson, P. (2007). Two forms of spoken repetition in a girl with autism. *International Journal of Language & Communication Disorders*, 42, 427-444.
- Suchman, L. (1997). Centres of coordination: A case and some themes. In Lauren B. Resnick, Roger Säljö, Clotilde Pontecorvo, Barbara Burge (Eds.), *Discourse, tools, and reasoning: Essays on situated cognition* (pp. 41-62), Berlin: Springer.
- Summit, S.A. (1978). The witness needs help. In American Bar Association, *The deposition: A simulation with commentary* (pp. 110-114). Chicago: American Bar Association.
- Suomi, S.J. (1997). Nonverbal communication in nonhuman primates: Implications for the emergence of culture. In U. Segerstråle & P. Molnár (Eds.), *Nonverbal communication: Where nature meets culture*. Mahwah, NJ: Lawrence Erlbaum.
- Sutherland, O. & Couture, S. (2007). The discursive performance of the alliance in family therapy: A conversation analytic perspective. *Australian and New Zealand Journal of Family Therapy*, 28(4), 210-217.
- Szymanski, M.H., Vinkhuyzen, E., Aoki, P., & Woodruff, A. (2006). Organizing a remote state of incipient talk: Push-to-talk mobile radio interaction. *Language in Society* 35, 393-418.
- Talsma, D., Doty, T.J., Strowd, R., & Woldorff, M.G. (2006). Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. *Psychophysiology*, 43, 541-549.
- Tannen, D. (1984). The pragmatics of cross-cultural communication. *Applied Linguistics*, 5, 189-195.
- Tannen, D. (1985). Silence: Anything but. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 93-111). Norwood, NJ: Ablex.
- ten Have, P. (1989). The consultation as a genre. In B. Torode (Ed.), *Text and talk as*

- social practice* (pp. 115-135). Dordrecht/Providence: Foris Publications.
- ten Have, P. (1990). Methodological issues in conversation analysis. *Bulletin de Méthodologie Sociologique*, 27, 23-51.
- This Wicked Day (2010, June 16) Capitalisation is a political act. [Web log comment]  
Retrieved from <http://wickedday.wordpress.com/2010/06/16/capitalisation-is-a-political-act/> (2011, February 12)
- Tomicic, A., & Martínez Guzmán, C. (2011). Voice and psychotherapy: Introduction to a line of research on mutual regulation in psychotherapeutic dialog. *Revista de Psicología*, 13(20), 109-139.
- Tomlinson, T.M., & Hart, J.T., Jr. (1962). Validation study of the progress scale. *Journal of Consulting Psychology*, 26, 74-78.
- Varney, S. (2012, March 22). *What do people actually do when watching situation comedy programmes together*. Paper presented at Discourse-Communication-Conversation, Loughborough, UK.
- Viney, R., & Berger, I. (2010, July 8). *That's what I'm talking about: How television plays into sequence and topic organisation*. Paper presented at the 10th International Conference on Conversation Analysis. Mannheim, Germany.
- Voilmy, D., & Ting, K.L.H. (2007). A first-time wireless internet connection: More than just clicking on a link. *PsychNology Journal*, 5(2), 165-195.
- vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic Interaction*, 24, 198-216.
- Walker, A., Rablin, A., & Rogers, C. (1960). Development of a scale to measure process change in psychotherapy. *Journal of Clinical Psychology*, 16, 79-85.
- Walker, A.G. (1985). The two faces of silence: The effect of witness hesitancy on lawyers' impressions. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (55-75). Norwood, NJ: Ablex.

- Walker, G. (2010). The phonetic constitution of a turn-holding practice: Rush-throughs in English talk-in-interaction. In D. Barth-Weingarten, E. Reber, & M. Selting, *Prosody in interaction* (pp. 51-72). Amsterdam: John Benjamins.
- Watzlawick, P., Beavin, J., & Jackson, D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York: Norton.
- West, C. (1979). Against our will: Male interruptions of females in cross-sex conversation. *Annals of the New York Academy of Sciences*, 327, 81-79.
- Whitehead, K.A. (2011). Some uses of head nods in “third position” in talk-in-interaction. *Gesture*, 11(2), 103-122.
- Wiggins, S. (2002). Talking with your mouth full: Gustatory 'mmm's and the embodiment of pleasure. *Research on Language and Social Interaction*. 35, 311-336.
- Wilson, M., & Wilson, T.P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, 12, 957-968.
- Wilson, T.P., and Zimmerman, D.H. (1986). The structure of silence between turns in two-party conversation. *Discourse Processes*, 9, 375-390.
- Woodruff, A., & Aoki, P. (2004a). Push-to-talk social talk. *Computer Supported Cooperative Work*, 13, 409-441.
- Woodruff, A., & Aoki, P. (2004b). Conversation analysis and the user experience. *Digital Creativity*, 15(4), 232-238.
- Yngve, V.H. (1970). On getting a word in edgewise. In Papers from the Sixth Regional Meeting of the Chicago Linguistic Society (pp. 567-577). Chicago: Chicago Linguistic Society
- Zimmerman, D. H., & West, C. (1975). Sex roles, interruptions and silences in conversation. In B. Thome & N. Henley (Eds.), *Language and sex: Difference and dominance* (pp. 105-129). Rowley, MA: Newbury House.

### Appendix A – Transcription Conventions

The transcription system used in Conversation Analysis was initially developed by Gail Jefferson and has been modified over time by various researchers to better represent particular phenomena. It aims to capture, graphically, the structure of talk as it emerges, so particular attention is paid to overlap and silence. Conversation analysts often use an informal phonetic method to capture the sound of talk when standard orthography would be misleading or ambiguous.

Symbol	Meaning
>>	line(s) most relevant to the in-text discussion
(( ))	transcriptionist note or non-vocal behaviour
(.)	silence less than 0.1 second
(1.5)	silence in seconds and tenths of sections; i.e. (1.5) is one and a half second of silence
°	quieter than surrounding speech
>	slower talk turning into faster talk
<	faster talk turning into slower talk
,	slight terminal rise in intonation
.	terminal drop in intonation
↑↓	sharp rise or fall, respectively, in intonation
:	stretched sound
[ ]	overlapping talk
( )	unclear speech
–	sound cut off
£	smile voice

=	when occurring at the end of a line and at the beginning of another line, indicates that the lines are ‘through produced’ as one turn when occurring anywhere else, indicates that there is no silence between words or syllables (i.e. ‘latched’ sounds)
hh	audible outbreath
.hh	audible inbreath
huhehhuhee	laughter with approximated syllables
ex (h) amphhle	laughter within words – (h) indicates laughter that is formed separately from the syllable, whilst h indicates laughter that is formed as part of the syllable



## 1. ETHICS BOARD

### ***RESEARCH PARTICIPANT CONSENT FORM***

#### **Title and brief description of Research Project:**

##### **Conversation Analytic Data Bank**

Thank-you for your considering taking in this research project. We are collecting video film data of people talking and undertaking everyday activities together in a variety of settings. This might include eating a meal with others, playing cards, washing up, decorating the house, watching television, dancing, having a game of golf or just having a chat – but it can include all sorts of other possibilities in which there is some talk or interaction between two or more people. We are interested in ordinary interaction as it is ordinarily done – so there is no requirement to behave any differently than you normally do. The research is simply interested in how we engage in social interaction – from studying people talking in everyday settings it is hoped that we can get a better understanding regarding the use of talk, gesture, gaze and such like in everyday interaction.

Once you have completed filming you will be invited to delete or identify for deletion any parts of the film that you wish (you can delete the entire film if you would like to). The film will be analyzed to explore features of the interaction so as to provide a better understanding about how people interact. Thus we might look at how gestures are used, how we co-ordinate gaze and talk, how movement is integrated into activity sequences or how intonation shapes are used and responded to – anything which helps deepen our understanding about how we interact with one another.

Please note, you can withdraw from the study or delete any part of the film in which you appear at any moment. If you have any remaining questions concerning the research that you have been involved in please contact either Paul Dickerson or John Rae using the contact details below.

#### **Name and status of Investigators:**

Dr Paul Dickerson ([p.dickerson@roehampton.ac.uk](mailto:p.dickerson@roehampton.ac.uk) 020 8392 3613)

Dr John Rae ([j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk) 020 8392 3612)

Estefania Guimaraes ([E.Guimaraes@roehampton.ac.uk](mailto:E.Guimaraes@roehampton.ac.uk) (020 8392 3000 ext 4515)

Please read the consent form and feel free to decline entirely or only agree to those aspects that you are completely happy with.

**PART 1. CONSENT FOR YOU TO BE FILMED FOR RESEARCH PURPOSES.**

The study involves a video recording being made of you. Some of the data recorded may involve interactions with the researcher(s). You can leave the scene of the recording or request that the filming stops at any stage.

You may withdraw your consent at any time before, during or after the filming. You can request the deletion of any part of the film in which you appear.

We would like to make transcripts of the talk and actions that happen in the video and store them for research purposes. They are likely to be of research value long into the future.

Are you happy to be filmed as part of this project?

yes / no (if "yes" please sign) \_\_\_\_\_

**Consent Statement:**

I agree to take part in this research, and am aware that I am free to withdraw at any point. I understand that the information I provide will be treated in confidence by the investigator although I understand that if a risk of serious harm arises that the researchers may need to take appropriate action. I understand that my identity will be protected in the publication of any findings unless I specifically choose for my recording, or stills from it, to be seen by other people in part 2 or part 3 below.

yes / no (if "yes" please sign) \_\_\_\_\_

Name: \_\_\_\_\_

Please note: if you have a concern about any aspect of your participation, please raise this with the investigator, or with the Dean of School who is

Name: Mr Michael Barham

Contact Address: School of Human and Life Sciences; Roehampton University, Whitelands College, Holybourne Avenue, LONDON, SW15 4JD

Direct Phone No: 020 8392 3620 Email: [m.barham@roehampton.ac.uk](mailto:m.barham@roehampton.ac.uk)



**PART 2. CONSENT FOR FILM AND SOUNDTRACK TO BE USED FOR RESEARCH PURPOSES.****Transcripts of your talk and actions**

2. Do you give permission for written transcripts of your talk and actions to be presented at conference & workshop presentations?

yes / no (if "yes" please sign) \_\_\_\_\_

3. Do you give permission for written transcripts of you talk and actions appear in academic and professional publications (such as journals, including electronic and paper versions)

yes / no (if "yes" please sign) \_\_\_\_\_

**Photo stills of you**

4. Do you give permission for still photographs of you to be presented in conference & workshop presentations?

yes / no (if "yes" please sign) \_\_\_\_\_

5. Do you give permission for still photographs of you to be presented in research and educational publications?

yes / no (if "yes" please sign) \_\_\_\_\_

**Audio fragments**

6. Do you give permission for audio fragments involving you to be presented in conference & workshop presentations?

yes / no (if "yes" please sign) \_\_\_\_\_

7. Do you give permission for audio fragments involving you to be presented in research and educational publications?

yes / no (if "yes" please sign) \_\_\_\_\_

**Video fragments**

8. Do you give permission for video fragments involving you to be presented in conference & workshop presentations?

yes / no (if "yes" please sign) \_\_\_\_\_

9. Do you give permission for video fragments involving you to be presented in research and educational publications?

yes / no (if "yes" please sign) \_\_\_\_\_

**PART 3. DATA TO BE SHARED WITH OTHER RESEARCHERS**

10 .Do you give permission for transcripts involving you to be shared with other researchers?

yes / no (if "yes" please sign) \_\_\_\_\_

11. Do you give permission for audio recordings involving you to be shared with other researchers?

yes / no (if "yes" please sign) \_\_\_\_\_

12. Do you give permission for photo stills involving you to be shared with other researchers?

yes / no (if "yes" please sign) \_\_\_\_\_

13. Do you give permission for video recordings involving you to be shared with other researchers?

yes / no ((if "yes" please sign) \_\_\_\_\_

Date: \_\_\_\_\_

**Appendix C – Final Good Relations Ethical Approval**

--- On Mon, 17/8/09, L.Rochard@roehampton.ac.uk <L.Rochard@roehampton.ac.uk> wrote:

From: L.Rochard@roehampton.ac.uk <L.Rochard@roehampton.ac.uk>

Subject: Ethics Application - BEREN, Israel

To: chezsruli@yahoo.com

Cc: P.Stribling@roehampton.ac.uk, j.rae@roehampton.ac.uk, L.Slade@roehampton.ac.uk,

Jan.Harrison@roehampton.ac.uk

Date: Monday, 17 August, 2009, 2:55 PM

Dear Israel ,

Ethics Application (Research Student)

Title: Good relations: practices that foster open communication between healthcare providers and TGI patients

Ref: PSY 09/028

Many thanks for sending in your response to the two conditions set by the Ethics Board. I am pleased to confirm that you have met these two conditions and your application has been approved. We do not require anything further in relation to your University ethics application.

I understand that you will be sending me a hard copy of your NHS ethics application form for the Chair of the Ethics Board to approve once you have finalised this.

Many thanks,

Lemady

Lemady Rochard

Research Policy Officer

Research and Business Development Office

208 Grove House, Froebel College

Roehampton University

Roehampton Lane

London

SW15 5PJ

T: +44 (0)20 8392 3256

E: L.Rochard@roehampton.ac.uk



Dear Service User,

I am Israel Berger, and I am researching patient-provider communication. The study is called Good Relations: Practices That Foster Open Communication Between Healthcare Providers and TGI Service Users. TGI is an acronym that encompasses transsexual, transgender, gender independent, and intersex. I am looking for non-TGI service users to participate in the study to compare interactions among different groups. This research will examine patient-provider interactions and potentially identify practices that work well in fostering open communication with TGI service users and compare how TGI and non-TGI service users respond to those practices. I hope to generate results that will be of use to the medical and mental healthcare community in dealing with TGI people as well as contribute to the development of further research in the areas of TGI health and patient-provider relationship. I am doing this research as part of my doctoral work at Roehampton University with supervisor Dr John Rae (+44 208 8392 3612, [j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk)).

If you are interested in participating, please contact me. To make things easier for you and your healthcare provider(s), there is a form at the end of this letter that I would kindly ask that you fill in and return to me. I can then contact your provider with the appropriate information. If you and your healthcare provider(s) decide to take part in this study here is what will happen: Your appointment(s) or interaction(s) with your healthcare provider(s) will be audio or video recorded depending on what you consent to and in a manner that is most convenient to you and your provider(s) who participate. Please note that it is only possible to participate if your healthcare provider(s) agree to participate as well. No discomfort or risks are expected from participating in this study aside from initial nervousness or uneasiness from knowing that you are being recorded. You may choose to have as many sessions as you feel comfortable being recorded whether that is one or many during the study period. You may withdraw at any time and without giving a reason.

Before participating, you will receive further information and a consent form. You will be able to ask any questions you may have of the researcher as well.

All records will be kept in the utmost confidence and published work will not identify you unless you choose to be identified (such as by allowing me to present recordings at academic or medical presentations). You will be able to choose how the recordings and transcripts are used and can change the permissions or restrictions for certain uses at any time or to withdraw completely from the study.

Thank you for your consideration. I look forward to hearing from you. I can be reached by telephone at (303) 800 6736 or +44 7588214820 or by e-mail at [Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk).

Regards,

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588214820

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
j.rae@roehampton.ac.uk  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
m.barham@roehampton.ac.uk  
+44 20 8392 3617

**Response - Good Relations Study**

Your title (e.g. Dr., Ms.):

Your name:

Your Preferred Pronoun (e.g. zie, she, he):

Your e-mail address:

Your daytime telephone number:

Your evening telephone number:

What is your preferred method of contact?:

Your correspondence address:

If there is anything additional that you would like me to know, please write it here.



Dear Service User,

I am Israel Berger, and I am researching patient-provider communication with transsexual, transgender, gender independent intersex (TGI) people as patients or service users. The study is called Good Relations: Practices That Foster Open Communication Between Healthcare Providers and TGI Service Users. I am looking for participants who are trans, of trans experience, have a non-binary gender or genders, and/or are intersex (or somewhere on the trans spectrum and/or intersex spectrum, I am not looking to include only particular identities or experiences). This research will examine patient-provider interactions and potentially identify practices that work well in fostering open communication with TGI service users. I hope to generate results that will be of use to the medical and mental healthcare community in dealing with TGI people as well as contribute to the development of further research in the area of TGI health. I am doing this research as part of my doctoral work at Roehampton University with supervisor Dr. John Rae (+44 208 8392 3612, [j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk)).

If you are interested in participating, please contact me. To make things easier for you and your healthcare provider(s), there is a form at the end of this letter that I would kindly ask that you fill in and return to me. I can then contact your provider with the appropriate information. If you and your healthcare provider(s) decide to take part in this study here is what will happen: Your appointment(s) or interaction(s) with your healthcare provider(s) will be audio or video recorded depending on what you consent to and in a manner that is most convenient to you and your provider(s) who participate. Please note that it is only possible to participate if your healthcare provider(s) agree to participate as well. No discomfort or risks are expected from participating in this study aside from initial nervousness or uneasiness from knowing that you are being recorded. You may choose to have as many sessions as you feel comfortable being recorded whether that is one or many during the study period. You may withdraw at any time and without giving a reason.

Before participating, you will receive further information and a consent form. You will be able to ask any questions you may have of the researcher as well.

All records will be kept in the utmost confidence and published work will not identify you unless you choose to be identified (such as by allowing me to present recordings at academic or medical presentations). You will be able to choose how the recordings and transcripts are used and can change the permissions or restrictions for certain uses at any time or to withdraw completely from the study.

Thank you for your consideration. I look forward to hearing from you. I can be reached by telephone at (303) 800 6736 or +44 7588214820 or by e-mail at [Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk).

Regards,

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588214820

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party

please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
j.rae@roehampton.ac.uk  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
m.barham@roehampton.ac.uk  
+44 20 8392 3617



**Response – Good Relations Study**

Your title (e.g. Dr., Ms.):

Your name:

Your preferred pronoun (e.g. zie, she, he):

Your e-mail address:

Your daytime telephone number:

Your evening telephone number:

What is your preferred method of contact?:

Your correspondence address:

If there is anything additional that you would like me to know, please write it here.



Dear [name],

I am Israel Berger, and I am working toward a doctoral degree at Roehampton University with supervisor Dr. John Rae (+44 208 8392 3612, [j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk)). I am researching patient-provider communication with transgender, transsexual, gender independent, and intersex (TGI) people as patients or service users. The study is called Good Relations: Practices That Foster Open Communication Between Healthcare Providers and TGI Service Users. This research will examine patient-provider interactions and potentially identify practices that work well in fostering open communication with TGI patients. Non-TGI patients or service users are welcome to participate in this study as well. These data will be used to compare how service users from different groups may react to similar practices. I hope to generate results that will be of use to the medical and mental healthcare community in dealing with TGI people as well as contribute to the development of further research in the areas of TGI health and patient-provider relationship.

If you are interested in participating, please contact me. To make our initial communication easier, there is a form at the end of this letter that I would kindly ask that you fill in and return to me. I will then be able to send you the appropriate information to pass to your patient(s). If you and your patient(s) or service user(s) decide to take part in this study here is what will happen: Your appointment(s) or interaction(s) with your patient(s) or service user(s) will be audio or video recorded depending on what you consent to and in a manner that is most convenient to you and the patient(s) or service user(s) who participate. Please note that it is only possible to participate if your patient(s) or service user(s) agree to participate as well. No discomfort or risks are expected from participating in this study aside from initial nervousness or uneasiness from knowing that you are being recorded. You may choose to have as many sessions as you feel comfortable being recorded whether that is one or many during the study period. You may withdraw at any time and without giving a reason.

Before participating, you will receive further information and a consent form. You will be able to ask any questions you may have of the researcher as well.

All records will be kept in the utmost confidence and published work will not identify you unless you choose to be identified (such as by allowing me to present recordings at academic or medical presentations). You will be able to choose how the recordings and transcripts are used and can change the permissions or restrictions for certain uses at any time or to withdraw completely from the study.

Thank you for your consideration. I look forward to hearing from you. I can be reached by telephone at (303) 800 6736 or +44 7588214820 or by e-mail at [Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk).

Regards,

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588214820

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party

please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
j.rae@roehampton.ac.uk  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
m.barham@roehampton.ac.uk  
+44 20 8392 3617

**Response – Good Relations Study**

Your title (e.g. Dr., Ms.):

Your name:

Your preferred pronoun (e.g. zie, she, he):

Your e-mail address:

Your daytime telephone number:

Your evening telephone number:

What is your preferred method of contact?:

Your correspondence address:

If there is anything additional that you would like me to know, please write it here.

**Title of Research Project**

Good Relations: Practices That Foster Open Communication  
Between Healthcare Providers and TGI Patients

**Investigator Contact Details**

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588214820

Thank you for considering taking part in this research project. The researcher will explain the project to you in detail. Please feel free to ask questions. If you have more questions later, Israel Berger, the person mainly responsible for this study, will discuss them with you. His contact information is as follows: (303) 800 6736 or +44 7588214820, [Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)

*Description of the project:*

This research will examine patient-provider interactions to identify practices that work well and less well in fostering open communication with transsexual, transgender, gender independent, and intersex (TGI) patients. Non-TGI patients or service users are welcome to participate in this study as well. These data will be used to compare how service users from different groups may react to similar practices. I hope to generate results that will be of use to the medical and mental healthcare community in dealing with TGI patients as well as contribute to the development of further research in the area of TGI health.

*What will be done:*

If you decide to take part in this study here is what will happen: Your appointment(s) or interaction(s) with your healthcare provider(s) or patient(s) will be audio or video recorded depending on what you consent to. You may participate for one or many appointments during the project. The project will last approximately three years, but you do not need to participate for the entire duration of the project.

*Risks or discomfort:*

No discomfort or risks are expected from participating in this study aside from initial nervousness or uneasiness from knowing that you are being recorded.

*Benefits of this study:*

There may not be a direct benefit to you. Participants may choose to have the researcher share observations regarding the interaction that will not necessarily be present in published works. This may lead to improved communication and an improved relationship between the participants or other healthcare situations such as with other providers or service users. A summary of the results of the study will be available. Your participation in the study will help the researcher to identify ways in which healthcare providers who work with TGI people can help to improve communication and construct mutually beneficial treatment outcomes. Ultimately this research is designed to identify specific ways in which the patient-provider relationship can be enhanced for TGI people and their providers.

*Confidentiality:*

Your part in this study is confidential unless you consent to identifying information (such as video clips that show your face) being used in publications and/or presentations. Options for you to indicate what levels of disclosure you will or will not allow will be presented in the consent form.

Your consent to use identifying information will be re-confirmed before any identifying information is published.

*Storage of Data:*

Original recordings will be kept in locked containers under the control of the researcher. Digital versions will be stored on the researcher's password-protected computer and/or on flash drives or discs in password protected files. Consent forms, raw data, and processed data (e.g., with pseudonyms) will be kept separately. Roehampton University requires that all records be kept for a minimum of six (6) years following the last publication.

*Analysis:*

Conversation analysis will be used to analyze the data from this project. This involves analysis of how interactions develop and turns at talk are managed. It is important to discuss ones findings in 'data sessions' where one shares ideas with other conversation analysts, have their feedback about ones analysis, and develop subsequent analyses and ideas. In this project, they will happen most with academic supervisors, if you agree to that use of the recordings. You may also consent to the recordings being shared with other researchers for this purpose. In either case, if you agree to recordings being played, you can opt to allow unedited versions (with identifying information) or versions in which identifying details have been erased. The data would only be shared with researchers, selected by Mr. Berger, who will and treat it with respect and with due care and attention and will use if for serious research purposes.

*Decision to withdraw at any time:*

The decision to take part in this study is up to you. You do not have to participate. If you decide to take part in the study, you may withdraw at any time and without giving reason. Whatever you decide will in no way penalize you. If you wish to withdraw, simply inform Israel Berger, (303) 800 6736 or +44 7588214820, Israel.Berger@roehampton.ac.uk, of your decision. Please notify him as well of any changes you wish to make regarding the permissions or restrictions you may have given for the publication or presentation of recordings or transcripts (see below).

If you consent to identifying information being published, you may also notify him of your decision to not have it published any longer. He will then remove them from any website or other media to which he has access. Please note that once something has been published, it could still be circulated even if publication is discontinued and that it is generally not possible to retract the content of books, journal articles, and other media to which the researcher does not have direct access.

If at any time during the study you become unable to consent for yourself, any data already collected will be retained and used in the study and further recordings will not be made. However, no identifying information will be published or shared once you are unable to give consent, even if you have already consented to these uses of the data. If you want this to be handled any other way, please let Israel Berger, know in writing (this can be written on the consent form).

*Publication:*

As part of this research you will have made recordings of some of your interactions. I would like you to indicate below to what uses of these recordings and their transcripts you are willing to consent. This is completely up to you. I will only use the recordings and transcripts in ways in which you agree. It cannot be guaranteed that the transcripts or recordings will not be misused by others once they are in the public domain. You should consider carefully whether you want identifying information to be used for websites or publications. Identifying information means anything that could identify you easily, such as your name, face, or uncommon details which could be your PCT or your profession if mentioned in the interaction. You may have some things that you don't mind other people to see/hear and other things you don't want people to see/hear. In addition to the list below, you can make notes next to an item if you are concerned about it in particular or in the space provided if you are concerned about something in general. You may choose to delete any portion of the recording in which you are recorded. Simply inform me of your decision.

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
j.rae@roehampton.ac.uk  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
m.barham@roehampton.ac.uk  
+44 20 8392 3617

**Title of Research Project**

Good Relations: Practices That Foster Open Communication  
Between Healthcare Providers and TGI Patients

**Investigator Contact Details**

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588214820

Please note that by consenting to the research, you agree to being AUDIO recorded and written transcripts being published, shared, or presented WITHOUT identifying information. This means that unless you consent otherwise, pseudonyms will be used for names, locations, and other identifying details will be changed and only the written transcript will be used in publication, sharing, or presentation. Your consent to use identifying information will be re-confirmed before any identifying information is published.

Please **initial** below to indicate those items to which you give additional consent.

1. I agree to being video recorded \_\_\_\_\_
2. I agree to the recordings being used for publications in the form of written transcripts WITH identifying information (for example, if you have an uncommon occupation and are discussing its effects on your health) \_\_\_\_\_
3. I agree to photo stills (if video recorded) of me being used for publications with written transcripts (This might be done to show where participants are in a room, for example.) \_\_\_\_\_
4. The recordings can be shared with research supervisors and examiners for degree requirements  
\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information
5. The recordings can be studied by other researchers selected by Mr. Berger  
\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information
6. The recordings can be played in professional, academic and educational contexts  
\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information
7. The written transcript can be kept in an archive for other researchers  
\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information
8. The recordings can be kept in an archive for other researchers.  
\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information
9. Brief extracts from the recordings can be placed on an academic website, DVD, or other media to support scholarly publications



\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information

10. Brief extracts from the recordings can be placed on an academic website, DVD, or other media for training healthcare providers

\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information

11. The recordings can be played in public presentations to non-scientific groups

\_\_\_\_\_WITHOUT identifying information \_\_\_\_\_WITH identifying information

Any restrictions or permissions in addition to those above:

---



---



---



---



---

I have read the information sheet and give my consent for the use of the recordings as indicated above. I agree to take part in this research, and am aware that I am free to withdraw at any point. I understand that the information I provide will be treated in confidence by the researcher and that my identity will be protected in the publication of any findings unless I give consent otherwise.

Name ..... Would you like to receive a summary of the results? Y / N

Signature ..... Would you like to receive personalized observations? Y / N

Date .....

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
j.rae@roehampton.ac.uk  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
m.barham@roehampton.ac.uk  
+44 20 8392 3617



Dear [name],

Thank you for participating in my study. I very much appreciate you taking the time to allow me to collect data for the Good Relations project. I would like to remind you that if you want to change the permissions or restrictions on my use of the recordings or transcripts or to withdraw completely, you may do so at any time. You have a copy of the consent form. If you want to add or remove anything with regard to the permissions or restrictions, please do not hesitate to e-mail me or send me a letter. If you would like to discuss or have any questions about the research, please feel free to call or e-mail me. My contact information can be found near the end of this letter.

*Description of the project and researcher:*

I am Israel Berger, and I am working toward a doctoral degree at Roehampton University with supervisor Dr. John Rae (+44 208 8392 3612, [j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk)). I am researching patient-provider communication with transgender, transsexual, gender independent, and intersex (TGI) people as patients or service users. The study is called Good Relations: Practices That Foster Open Communication Between Healthcare Providers and TGI Service Users. This research will examine patient-provider interactions and potentially identify practices that work well in fostering open communication with TGI service users. Non-TGI patients or service users are welcome to participate in this study as well. These data will be used to compare how service users from different groups may react to similar practices. I hope to generate results that will be of use to the medical and mental healthcare community in dealing with TGI patients as well as contribute to the development of further research in the area of TGI health.

Analysis will detail verbal and non verbal communication features. You may choose to have as many sessions as you feel comfortable being recorded whether that is one or many during the study period. You may withdraw at any time and without giving a reason.

There may not be a direct benefit to you. Participants may choose to have the researcher share observations regarding the interaction that will not necessarily be present in published works. This may lead to improved communication and an improved relationship between the participants or other healthcare situations such as with other providers or service users. A summary of the results of the study will be available. Your participation in the study will help me to identify ways in which healthcare providers who work with TGI people can help to improve communication and construct mutually beneficial treatment outcomes. Ultimately this research is designed to identify specific ways in which the patient-provider relationship can be enhanced for TGI people and their providers.

*Decision to withdraw at any time:*

The decision to take part in this study is up to you. You do not have to participate. If you decide to take part in the study, you may withdraw at any time and without giving reason. Whatever you decide will in no way penalize you. If you wish to withdraw, simply inform Israel Berger, (303) 800 6736 or +44 7588 214820, [Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk), of your decision. Please notify him as well of any changes you wish to make regarding the permissions or restrictions you may have given for the publication or presentation of recordings or transcripts (see below).

If you consent to identifying information being published, you may also notify him of your decision to not have it published any longer. He will then remove them from any website or other media to which he has access. Please note that once something has been published, it could still be circulated even if publication is discontinued and that it is generally not possible to retract the content of books, journal articles, and other media to which the researcher does not have direct access.

If at any time during the study you become unable to consent for yourself, any data already collected will be retained and used in the study and further recordings will not be made. However,

no identifying information will be published or shared once you are unable to give consent, even if you have already consented to these uses of the data. If you want this to be handled any other way, please let Israel Berger, know in writing (this can be written on the consent form).

*Publication:*

As part of this research you will have made recordings of some of your interactions. I will only use the recordings and transcripts in ways in which you agree. It cannot be guaranteed that the transcripts or recordings will not be misused by others once they are in the public domain. You should consider carefully whether you want identifying information to be used for websites or publications. Identifying information means anything that could identify you easily, such as your name, face, or uncommon details which could be your PCT or your profession if mentioned in the interaction. You may have some things that you don't mind other people to see/hear and other things you don't want people to see/hear. In addition to the list below, you can make notes next to an item if you are concerned about it in particular or in the space provided if you are concerned about something in general. You may choose to delete any portion of the recording in which you are recorded. Simply inform me of your decision.

Please note that by consenting to the research, you agree to being AUDIO recorded and written transcripts being published, shared, or presented WITHOUT identifying information. This means that unless you consent otherwise, pseudonyms will be used for names, locations, and other identifying details will be changed and only the written transcript will be used in publication, sharing, or presentation.

I have included with this letter a list of organizations and support groups that deal with TGI issues and in particular healthcare. Thank you again for your time and participation.

Regards,

Mr. Israel Berger  
School of Human & Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[Israel.Berger@roehampton.ac.uk](mailto:Israel.Berger@roehampton.ac.uk)  
(303) 800 6736 or +44 7588 214820

Please note: if you have a concern about any aspect of your participation or any other queries please raise this with the researcher. However if you would like to contact an independent party please contact the Dean of School (or if the researcher is a student you can also contact the Director of Studies.)

**Director of Studies Contact Details:**

Dr. John Rae  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[j.rae@roehampton.ac.uk](mailto:j.rae@roehampton.ac.uk)  
+44 20 8392 3612

**Dean of School Contact Details:**

Mr. Michael Barham  
School of Human and Life Sciences  
Whitelands College  
Holybourne Avenue  
London SW15 4JD  
[m.barham@roehampton.ac.uk](mailto:m.barham@roehampton.ac.uk)  
+44 20 8392 3617

*TGI Organizations and Support Groups:*

**Organization Intersex International (Organization Internationale des Intersexués)**

[www.intersexualite.org](http://www.intersexualite.org)

Please see **TG Forum's** Resource Database for local, national, and international organizations and support groups. <http://www.tgforum.com/resourcesearch.php>

**The Gender Identity Research and Education Society**

[www.gires.org.uk](http://www.gires.org.uk)

Melverley  
The Warren  
Ashted, Surrey  
KT21 2SP, UK

Tel: +44 1372 801554  
Fax: +44 1372 272297

## **Information about Video Recording Research**

**Name of Researcher: Rowena Viney**

### **About this research**

I am a PhD student at the University of York and I am carrying out research on how people construct social relationships in everyday talk, particularly in relation to lesbian identity.

This research will contribute to work within the field of conversation analysis, which seeks to understand how people communicate with each other in everyday settings. This includes not only what is said but also the small details of interaction, such as how people use gaze and gesture. Therefore, such conversations are video-recorded to capture how people communicate non-verbally, for example through nodding or smiling.

### **How it works**

I would like you to video-record approximately 2 hours of ordinary interactions between you and your friends or family members. This could include, for example, everyday interactions such as mealtimes or having coffee, or planned events such as dinner parties. You will be given a video camera to make the recording, and you decide which interactions you record. If you change your mind about any of the recordings that you have made, you can erase them before returning the equipment to me.

### **What happens next**

I will make written transcripts of the recordings and I will maintain confidentiality by anonymising any names and places mentioned in the talk. Your personal details will not be passed to any other people. I would like to be able to show the video to other researchers and to publish clips from it, and because communication relies in part on facial expressions, I would like to be able to show your face. If you agree to allow me to show the recordings to other researchers, I can bleep out any names and places mentioned if you wish.

It would be helpful if researchers other than myself could be allowed to use the recordings in their own research into language and communication, as there is so much detail in interaction that can be studied. I will only share it in this way if you agree.

**Thank you for taking part.**

**If you have questions or concerns about this research please feel free to contact me:**

Rowena Viney  
Department of Sociology  
University of York  
Heslington  
York YO10 5DD  
Tel: 07968 442194 Email: raev500@york.ac.uk

Department of **Sociology**

## Video Recording Release Consent Form

This consent form **must be signed by everyone** who has participated in the recording.

I agree to the use of my video-recording by Rowena Viney for research purposes.

I understand that:

- written transcripts will be made of the recordings and that these may be included in written or oral reports of the study;
- the transcripts will be entirely anonymous (i.e. my name and other identifying features will not be included in any written or oral report of this study);
- the recordings and transcripts will not be used in any other way unless I agree.

The following list shows any additional ways in which I am willing for the recordings and transcripts to be used:

1. Stills of video-clips can be used by Rowena Viney in academic publications.  
Yes [ ] No [ ]
2. Video-clips can be shown by Rowena Viney in academic contexts (e.g. teaching) and professional meetings (e.g. conferences).  
Yes [ ] No [ ]
3. Extracts from transcripts and associated video-clips can be displayed by Rowena Viney on academic web sites in conjunction with publications.  
Yes [ ] No [ ]
4. Please “bleep out” my name and any other identifying features used in the recordings when playing publicly.  
Yes [ ] No [ ]
5. The recordings can be placed in an archive for other researchers to use.  
Yes [ ] No [ ]

**Note: all participants in the interactions you record must sign this form.** If there are more people taking part in this recording than the space below allows for your names, signatures and dates, please continue overleaf. Thank you.

Name: \_\_\_\_\_  
\_\_\_\_\_

Name:

Signature: \_\_\_\_\_  
\_\_\_\_\_

Signature:

Date: \_\_\_\_\_  
\_\_\_\_\_

Date:

Name: \_\_\_\_\_  
\_\_\_\_\_

Name:

Signature: \_\_\_\_\_  
\_\_\_\_\_

Signature:

Date: \_\_\_\_\_  
\_\_\_\_\_

Date:

## Permission to use extracts

Thursday, 20 December, 2012 9:45

From:

"Rowena Viney"

To:

"Israel Berger"

Dear Israel,

This is to confirm that you have permission to use extracts from RV-DN and RV-KB in your thesis. This is consistent with the participants' consent, that researchers other than myself can use their recordings.

Best wishes,  
Rowena

### Bibliography of Content Analysis Sample

- Adkins, B.A., & Nasarczyk, J. (2009). Asynchronicity and the ‘time envelope’ of online annotation: The case of the photosharing website, Flickr. *Australian Journal of Communication*, 35(3), 115-140.
- Aoki, P.M., Szymanski, M.H., & Woodruff, A. (2006, November). *Media spaces and the mobile world*. Paper presented at ACM Conference on Computer Supported Cooperative Work. Banff, Canada. Retrieved from <http://www.paulaoki.com/papers/cscw06-media.pdf>
- NOTE: Although this is an earlier version of Aoki et al. (2009), it is substantially different and was therefore included in the sample.
- Aoki, P.M., Szymanski, M.H., & Woodruff, A. (2009). Bringing media spaces back to the streets: Notes on the interplay of research on media space and mobile communication. In S. Harrison (Ed.), *Media space 20+ years of mediated life* (pp. 287-301). London: Springer.
- Arminen, I., & Weilenmann, A. (2009). Mobile presence and intimacy—Reshaping social actions in mobile contextual configuration. *Journal of Pragmatics*, 41, 1905-1923.
- Barelds, L. (2011). *Beurtwisseling en sequentieorganisatie in Facebook-interacties* [Turn-taking and sequence organisation in Facebook interactions] (Master’s thesis). Retrieved from [http://scripties.let.eldoc.ub.rug.nl/FILES/root/Master/DoorstroomMasters/Communicatie-enInfor/2011/BareldsA.R./MA\\_1486128\\_AR\\_Barelds.pdf](http://scripties.let.eldoc.ub.rug.nl/FILES/root/Master/DoorstroomMasters/Communicatie-enInfor/2011/BareldsA.R./MA_1486128_AR_Barelds.pdf)
- Bischoff, F.A. (2005). *San tsu shing explicated: The classical introduction to Chinese couplets I to XI*. Vienna: Austrian Academy of Sciences Press.
- Bovet, A., & Sormani, P. (2004). Order in disorder: Rule-governed actions as sociological phenomena. Paper presented at Sixth International Conference on Social Science



- Methodology: Amsterdam. Retrieved from  
[http://konference.fdvinfo.net/rc33/2004/Data/PDF/stream\\_02-05.pdf](http://konference.fdvinfo.net/rc33/2004/Data/PDF/stream_02-05.pdf)
- Clark, S.J. (2009). *Getting personal: Talking the psychotherapy session into being* (Doctoral dissertation). Retrieved from <http://aiemca.net/wp-content/uploads/2010/03/clark-thesis-talking-psychotherapy-into-being1.pdf>
- Clift, R. (2001). Meaning in interaction: The case of actually. *Language*, 77, 245-291.
- Couper-Kuhlen, E. (2010). Commentary on Stivers and Rossano: Mobilizing response. *Research on Language & Social Interaction*, 43, 32-37.
- Davies, A. (1989). Communicative competence as language use. *Applied Linguistics*, 10(2), 157-170.
- Dean, J.W. (2001). *The Rehnquist choice: The untold story of the Nixon appointment that redefined the Supreme Court*. New York: The Free Press.
- Elliston, B. (n.d.). *Sat-nav talk and in-car talk: A study of groups' linguistic engagement with the content of in-car satellite navigation audio* (Proposal). Retrieved from [http://scholar.googleusercontent.com/scholar?q=cache:kfOqGyBvVGcJ:scholar.google.com/&hl=en&as\\_sdt=0,5](http://scholar.googleusercontent.com/scholar?q=cache:kfOqGyBvVGcJ:scholar.google.com/&hl=en&as_sdt=0,5)
- Eriksson, M. (1995). Berättande och konflikt i ungdomars samtal: En analys [Storytelling and conflict in young people's conversation]. *Folkmålsstudier*, 36, S139-S163. Retrieved from [http://fnf.fi/Site/Widget/Editor/822/files/Eriksson\\_fms36.pdf](http://fnf.fi/Site/Widget/Editor/822/files/Eriksson_fms36.pdf)
- Fele, G. (2008). The collaborative production of responses and dispatching on the radio: Video analysis in a medical emergency call center. *Forum: Qualitative Social Research*, 9(2), Article 40.
- Ford, C.E. (2010). Questioning in meetings: Participation and positioning. In A.F. Freed & S. Ehrlich (Eds.), *"Why do you ask?" The function of questions in institutional discourse* (pp. 211-234). Oxford: Oxford University Press.

- Gardner, R. (2010). Question and answer sequences in Garrwa talk. *Australian Journal of Linguistics*, 30, 423-445.
- Gardner, R., Fitzgerald, R., & Mushin, I. (2009). The underlying orderliness in turn-taking: Examples from Australian talk. *Australian Journal of Communication*, 36(3), 65-90.
- Georgakopoulou, A. (2008). 'On MSN with buff boys': Self- and other-identity claims in the context of small stories. *Journal of Sociolinguistics*, 12, 597-626.
- Goldberg, J.A. (2004). The amplitude shift mechanism in conversational closing sequences. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 257-297). Amsterdam: John Benjamins.
- Goodwin, C. (2006). Retrospective and prospective orientation in the construction of argumentative moves. *Text and Talk*, 26, 441-460.
- Heath, C.C., & Luff, P., & Sellen, A.J. (1997). Reconfiguring media space: Supporting collaborative work. In K. Finn, A.J. Sellen, & S.B. Wilbur (Eds.), *Video-mediated communication* (pp. 232-347). Mahwah, NJ: Lawrence Erlbaum.
- Heath, C.C., & Luff, P. (1993). Disembodied conduct: Communication through video in a multi-media office environment. In G. Button (Ed.), *Technology in working order: Studies of work, interaction, and technology* (pp. 35-54). London: Routledge.
- Heath, C., & Luff, P. (2000). *Technology in action*. Cambridge: Cambridge University Press.
- Jones, R., & Thornborrow, J. (2004). Floors, talk and the organization of classroom activities. *Language in Society*, 33, 399-423.
- Katz, J.E., & Aakhus, M. (2002). Part III: Public performance: Social groups and structures. In J.E. Katz & M. Aakhus (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 223-225). Cambridge: Cambridge University Press.
- Kurzon, D. (2007a). Towards a typology of silence. *Journal of Pragmatics*, 39, 1673-1688.

- Kurzon, D. (2007b). Peters Edition v. Batt : The intertextuality of silence. *International Journal for the Semiotics of Law*, 20, 285-303.
- Kurzon, D. (2011). Moment of silence: Constitutional transparency and judicial control. *International Journal for the Semiotics of Law*, 24, 195-209.
- Kushida, S. (2011). Confirming understanding and acknowledging assistance: Managing trouble responsibility in response to understanding check in Japanese talk-in-interaction. *Journal of Pragmatics*, 43, 2716-2739.
- Lim, N.E. (2010). Hedging in search of solidarity: The role of Wo Jue De in women's language. Paper presented at First International Symposium on Chinese Language and Discourse, Los Angeles. Retrieved from <http://limnieng.bol.ucla.edu/Gender%20Linguistics%20Paper.pdf>
- Lynch, M. (2000). Response to Wes Sharrock. *Text*, 20, 541-544.
- McColgan, G. (2005). A place to sit. *Journal of Contemporary Ethnography*, 34, 410-433.
- McIlvenny, P. (1991). Some thoughts on the study of sign language talk. In K. Sajavaara, D. Marsh, & T. Keto, *Communication and discourse across cultures and languages* (pp. 187-201). Tampere: Finnish Association for Applied Linguistics.
- McIlvenny, P. (2011). Video interventions in "everyday life": Semiotic and spatial practices of embedded video as a therapeutic tool in reality TV parenting programmes. *Social Semiotics*, 21, 259-288.
- Mondada, L. (2008). Using video for a sequential and multimodal analysis of social interaction: Videotaping institutional telephone calls. *Forum: Qualitative Social Research*, 9(3), Article 39.
- Moore, R.J., Ducheneaut, N., & Nickell, E. (2007). Doing virtually nothing: Awareness and accountability in massively multiplayer online worlds. *Computer Supported Cooperative Work*, 16, 265-305.

- Olsher, D. (2004). Talk and gesture: The embodied completion of sequential actions in spoken interaction. In R. Gardner and J. Wagner (Eds.), *Second language conversations* (pp. 221-245). New York: Continuum.
- Pallotti, G. (2001). External appropriations as a strategy for participating in intercultural multi-party conversations. In A. Di Luzio, S. Günthner, & F. Orletti, *Culture in communication: Analyses of intercultural situations* (pp. 295-333). Amsterdam: John Benjamins.
- Peräkylä, A. (2011). After interpretation: Third-position utterances in psychoanalysis. *Research on Language & Social Interaction*, 44, 288-316.
- Raudaskoski, P. (2011). When lives meet live: Categorization work in a reality TV show and “experience work” in two home audiences. *Text & Talk*, 31, 619-641.
- Raymond, G. (2000). The voice of authority: The local accomplishment of authoritative discourse in live news broadcasts. *Discourse Studies*, 2, 354-379.
- Reinsch, N.L., Turner, J.W., & Tinsley, C.H. (2008). Multicommunicating: A practice whose time has come?. *The Academy of Management Review*, 33, 391-403.
- Relieu, M. (2005, April). Drawing and handwriting on mobile phones: Some conversational features of a pen-stroke IM client. In *Proceedings Conference on Seeing, Understanding, Learning in the Mobile Age*. Budapest. Retrieved from [http://www.fil.hu/mobil/2005/Relieu\\_final.pdf](http://www.fil.hu/mobil/2005/Relieu_final.pdf).
- Relieu, M. (2006). Remarques sur l’analyse conversationnelle et les technologies médiatisées. *Revue française de linguistique appliquée*, 11, 17-32.
- Relieu, M. (2009). Mobile phone ‘work’: Disengaging and engaging mobile phone activities with concurrent activities. In R.S. Ling & S.W. Campbell, *The reconstruction of space and time: Mobile communication practices* (pp. 215-230). New Brunswick, NJ: Transaction.

- Relieu, M., & Olszewska, B. (2004). La matérialisation de l'Internet dans l'espace domestique: Une approche située de la vie domestique. *Réseaux*, 123, 119-148.
- Robinson, J.D. (1998). Getting down to business: Talk, gaze, and body orientation during openings of doctor-patient consultations. *Health Communication Research*, 25(1), 97-123.
- Saigo, H. (2011). *The Japanese sentence-final particles in talk-in-interaction*. Amsterdam: John Benjamins.
- Schegloff, E.A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. Thompson (Eds.), *Interaction and grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E.A. (2002). Beginnings in the telephone. In J.E. Katz & M. Aakhus (Eds.), *Perpetual contact: Mobile communication, private talk, public performance* (pp. 284-300). Cambridge: Cambridge University Press.
- Schegloff, E.A. (2004) Answering the phone. In G.H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 63-107). Amsterdam: John Benjamins.
- Schegloff, E.A. (2005). On Integrity in inquiry...of the investigated, not the investigator. *Discourse Studies*, 7, 455-80.
- Schegloff, E. A. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N.J. Enfield & S.C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 70-96). London: Berg.
- Schegloff, E.A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Schegloff, E.A. (2010). Commentary on Stivers & Rossano: "Mobilizing response". *Research on Language and Social Interaction*, 43, 38-48.
- Schegloff, E.A. (2011). Word repeats as unit ends. *Discourse Studies*, 13, 367-380.

- Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8(4), 289-327.
- Sinwongsuwat, K. (2007). *Partials in English conversation: A study of grammar in interaction* (Doctoral dissertation). University of Wisconsin, Madison, WI.
- Skårup, T. (2004). Brokering and membership in a multilingual community of practice. In R. Gardner & J. Wagner, *Second language conversations* (pp. 40-57). London: Continuum.
- Soares de Carvalho, E.N. (2007). *Interação entre pares na educação infantil: Exclusão-inclusão de crianças com deficiência intelectual* [Interaction between peers in early childhood education: Inclusion-exclusion of children with intellectual disabilities] (Doctoral dissertation). Retrieved from [http://repositorio.bce.unb.br/bitstream/10482/1844/1/Tese\\_EreniceNataliaSCarvalho.pdf](http://repositorio.bce.unb.br/bitstream/10482/1844/1/Tese_EreniceNataliaSCarvalho.pdf)
- Stivers, T., & Rossano, F. (2010). A scalar view of response relevance. *Research on Language and Social Interaction*, 43, 49-56.
- Suchman, L. (1997). Centres of coordination: A case and some themes. In L.B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools, and reasoning: Essays on situated cognition* (pp. 41-62), Berlin: Springer.
- Szymanski, M.H. (2002). Producing text through talk: Question-answering activity in classroom peer groups. *Linguistics and Education*, 13, 533-563.
- Szymanski, M.H., Aoki, P., Grinter, R.E., Hurst, A., Thomson, J.D., & Woodruff, A. (2008). Sotto voce: Facilitating social learning in a historic house. *Computer Supported Cooperative Work*, 17, 5-34.
- Szymanski, M. H., Vinkhuyzen, E., Aoki, P., & Woodruff, A. (2006). Organizing a remote state of incipient talk: Push-to-talk mobile radio interaction. *Language in Society* 35, 393-418.

- Toye, M. (2006). *Interactional sociolinguistics beyond (the) face-to-face: An analysis of improvised television interaction and internet message board discourse* (Doctoral dissertation). Georgetown University, Washington, DC.
- Traverso, V. (1996). *La conversation familière: Analyse pragmatique des interactions*. Lyon: Presses Universitaires de Lyon.
- Voilmy, D., & Ting, K.L.H. (2007). A first-time wireless internet connection: More than just clicking on a link. *PsychNology Journal*, 5(2), 165-195.
- Walker, R. (2008). 'There was nothing to say and nobody said it': Silence, disconnection and interruptions of Gertrude Stein's writing voice during World War II. *Cultural Studies Review*, 14, 147-160.
- Woodruff, A., & Aoki, P. (2004a). Push-to-talk social talk. *Computer Supported Cooperative Work*, 13, 409-441.
- Woodruff, A., & Aoki, P. (2004b). Conversation analysis and the user experience. *Digital Creativity*, 15, 232-238.

[CA 052109 Dinner 1]

01 Cor: Was your steak tender enough?

02 (2.3)

03 Cor: huh?

04 Al ((single head shake)) >Cooked it too fast<

05 (0.8)

06 Cor: I cooked it too fa:st

07 (0.7)

08 Al: ((nodding)) >Y’cooked it too fast<

09 (0.8)

10 Cor: Really:

11 Al: ((nods several times))

12 Cor: I[had it down on lo:w

13 Al: [((gaze to Coral))

14 (3.7)

15 Al: (I [dunno)

16 Cor: [Seerisly

17 (1.7)

18 Cor: It was #on lo::w # (# nodding during speech)

19 (2.2)

20 Cor: ((pointing to kitchen)) That notch before low

21 (4.4)

22 Cor: Musta been a bad cut,

23 (10.2) ((Al eats steak during silence, Coral keeps

24 gaze on Al))